

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:48:35 ; Search time 67.8241 Seconds  
(without alignments)  
2628.807 Million cell updates/sec

Title: US-09-596-774-6  
Perfect score: 2441  
Sequence: 1 MAWWTLLFLMAAKVPRQI.....LSTATKOTYDALHMOTLAPR 461

Scoring table:  
BLASTSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues  
Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Listing first 45 summaries

Database :  
1: Genesegq16Dec04:\*  
2: Genesegq1980s:\*  
3: Genesegq2000s:\*  
4: Genesegq2001s:\*  
5: Genesegq2002s:\*  
6: Genesegq2003as:\*  
7: Genesegq2003bs:\*  
8: Genesegq2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2441	100.0	461	2	AAW85508 Leader-sc
2	1675	68.6	483	2	AAW82315 Chimeric
3	1526	62.5	443	2	AAW24025 Single ch
4	1503.5	61.6	428	2	AAW24027 Single ch
5	1308	53.6	615	2	AAW95056 GAL4-DT-s
6	1308	53.6	617	2	AAW95057 GAL4-DT-s
7	1304	53.4	637	2	AAW26982 (FRP5) -BT
8	1303	53.4	711	2	AAW26980 Fv(FRP5) -
9	1302	53.3	530	2	AAW95053 scFv(FRP5
10	1302	53.3	651	2	AAW05136 scFv(FRP5
11	1302	53.3	699	2	AAW05138 scFv(FRP5
12	1302	53.3	892	2	AAW05140 scFv(225
13	1302	53.3	892	2	AAW05143 scFv(2
14	1302	53.3	892	2	AAW05139 scFv(2
15	1302	53.3	895	2	AAW05142 scFv(FRP
16	1302	53.3	899	2	AAW05144 scFv(FRP
17	1302	53.3	1020	2	AAW05141 scFv(2
18	1284	52.6	240	2	AAW85494 ErbB-2-sp
19	1284	52.6	240	2	AAW05134 Single ch
20	1275	52.2	241	2	AAW26979 Fv heavy
21	1249	51.2	631	7	ADH34566 CE7-speci
22	1015.5	41.6	443	6	ABG74488 Humanised
23	1015.5	41.6	443	6	ABG744240 Chimeric
24	1013	41.5	242	4	AAW20433 Anti-Fix/
25	1011.5	41.4	473	2	AAW26646 Chimeric

## ALIGNMENTS

26	1011.5	41.4	514	2	AAW26647 Chimeric
27	981.5	40.2	250	2	AAW02278 Anti-c-er
28	981.5	40.2	250	2	AAW29261 Single ch
29	981.5	40.2	250	2	AAW22400 Single ch
30	981.5	40.2	250	2	AAW47012 Single ch
31	981.5	40.2	250	2	AAW53168 Anti-c-er
32	981.5	40.2	250	2	AAW80422 Anti-c-er
33	981.5	40.2	250	7	AAW00714 741F8 sFv
34	978.5	40.1	259	7	ADG32359 Precursor
35	975.5	40.0	259	7	ADG32322 Mouse scF
36	963.5	39.5	329	7	ADG32361 Precursor
37	963.5	39.5	667	2	AAW39573 Sequence
38	962.5	39.4	651	2	AAW26649 Chimeric
39	962.5	39.4	692	2	AAW26650 Chimeric
40	960.5	39.3	259	7	ADG32324 Mouse scF
41	930	38.1	444	2	AAW36845 Single ch
42	921	37.7	240	8	ADN16154 Mouse C7
43	920	37.7	320	7	ADG32358 Precursor
44	920	37.7	569	7	ADG32357 Precursor
45	920	37.7	618	7	ADG32339 Fusion pr

  

RESULT 1	AAW85508	standard; protein; 461 AA.
ID	AAW85508	(first entry)
AC	AAW85508;	
XX	16-MAR-1996	
DT	Leader-scFv(FRP5):lyc-2 hinge;zeta.	
XX		
DE	Single chain antibody; scFv; antibody engineering; antitumour;	
XX	tumour antigen binding; T-cell receptor; cytotoxic T-lymphocyte;	
KW	monoclonal antibody; erbB-2; cancer; cell targeting;	
KW	adoptive immunotherapy.	
XX		
OS	Synthetic.	
XX		
FH	Key	Location/Qualifiers
FT	Peptide	40..93
FT	Domain	/label= IGH_chain_leader
FT		94..819
FT		/label= scFv(FRP)5
FT	Peptide	261..322
FT		/label= lyc-2_hinge
FT	Domain	323..461
FT		/label= zeta_chain
XX		
XX	WO9530014-A1.	
XX	09-NOV-1995.	
XX	20-APR-1995;	95WO-EP001494.
XX	02-MAY-1994;	94EP-00810244.
XX	(CIBA ) CIBA GEIGY AG.	
XX	Groner B, Moritz D;	
XX	WPI, 1995-393085/50.	
XX	N-PSDB; AAT05783.	
XX		
XX	New bifunctional proteins for use in killing tumour cells - contg. a	
XX	tumour antigen binding domain, a hinge region and a zeta chain derived	
XX	from a T-cell antigen receptor.	
XX		
XX	Example 1; Page 29-31; 46pp; English.	

CC A bifunctional protein (AA85505) consists of a single chain antibody, scFv(FRP)5, directed against the tumour erbB-2 antigen, a hinge region, and a functional zeta chain obt. from a T-cell receptor. The protein is expressed in host cells, esp. cytotoxic T-lymphocytes, providing them with a defined tumour cell specificity enabling targeting to defined tumour cells and MHC-unrestricted and MHC-independent tumour destruction in vitro or in vivo

XX Sequence 461 AA;

Query Match 100.0%; Score 2441; DB 2; Length 461;  
Best Local Similarity 100.0%; Pred. No. 1.4e-160;  
Matches 461; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MAMWTLFLMAAKPKQIQVQSGPELKKPGEYTKISCKASGYFTYGMWVQAQAG 60  
DB 1 MAMWTLFLMAAKPKQIQVQSGPELKKPGEYTKISCKASGYFTYGMWVQAQAG 60  
QY 61 OGKMMGWINTSGESTFADDFKGRFDSLETSANTAYQINNLKSEDMATYFCARMEY 120  
DB 61 OGKMMGWINTSGESTFADDFKGRFDSLETSANTAYQINNLKSEDMATYFCARMEY 120  
QY 121 HGVPYWGQGTIVTVSSGGSGSGGSDIQLTQSHKFLSTSVGDRVSIITKASQD 180  
DB 121 HGVPYWGQGTIVTVSSGGSGSGGSDIQLTQSHKFLSTSVGDRVSIITKASQD 180  
QY 181 VNNAVAMVQKQKPGSKKLITYSASSRYTGVPSRFSGSGGPFTFTISSVQAEALAVYC 240  
DB 181 VNNAVAMVQKQKPGSKKLITYSASSRYTGVPSRFSGSGGPFTFTISSVQAEALAVYC 240  
QY 241 QGHFRPTFGSGTKLEIKALEISNSVMYFSSVVPVLQKVNSTTKPVLRTSPVHPGT 300  
DB 241 QGHFRPTFGSGTKLEIKALEISNSVMYFSSVVPVLQKVNSTTKPVLRTSPVHPGT 300  
QY 301 SQPQRPEDCRPGSVKGTGLDFLEDPKCYLLDGLIFTYGVITLALYRAKFSRSAAETPA 360  
DB 301 SQPQRPEDCRPGSVKGTGLDFLEDPKCYLLDGLIFTYGVITLALYRAKFSRSAAETPA 360  
QY 361 NLQDPNQLYNELNGRREYDYLEKKRADPDMGKQRRRPPQGVNALQKDMAEAY 420  
DB 361 NLQDPNQLYNELNGRREYDYLEKKRADPDMGKQRRRPPQGVNALQKDMAEAY 420  
QY 421 SEITGERRRGGKHDGLYQGLSTATKDTYDALHMOTLAPR 461  
DB 421 SEITGERRRGGKHDGLYQGLSTATKDTYDALHMOTLAPR 461

RESULT 2

AAW82315  
ID AAW82315 standard; protein; 483 AA.

XX AC AAW82315;  
XX DT 26-FEB-1999 (first entry)  
XX DE Chimeric CD19/CTCR protein.

XX KW Single chain; antibody; B-cell marker; CD19; scFvCD19; gene therapy;  
XX KW retroviral vector; tumour-associated antigen; cancer; immunoglobulin;  
XX KW CDB alpha chain; hinge region; theta chain; transmembrane domain;  
XX KW intracellular domain; T-cell receptor-CD3 complex.

XX OS Homo sapiens.  
XX OS Synthetic.

XX PN DE19720152-A1.

XX PD 05-NOV-1998.

XX PF 02-MAY-1997; 97DE-01020152.

XX PR 02-MAY-1997; 97DE-01020152.

PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX Schroeder A, Schwenkenbecher J, Bargou R;

XX WPI; 1998-584686/50.

DR N-PSDB; AAV73324.

XX Retroviral vector system for gene therapy - by expression of therapeutic gene in T cells.

XX Claim 9; Page 6; 10pp; German.

CC This sequence represents a novel chimeric CD19/CTCR protein which is composed of an immunoglobulin leader sequence, an anti-CD19 single chain antibody fragment (scFvCD19), a hinge region from the CDB alpha-chain and the transmembrane and intracellular domain of the T-cell receptor-CD3 complex theta chain. This sequence is used in a retroviral vector system for gene transfer and expression in primary human T cells which involves a retroviral vector containing cDNA corresponding to a therapeutic gene, a promoter selected from the SV40 immediate early promoter or the human IL-2 or MHC I promoter and a selectable marker, and a packaging cell line (i.e. a modified retroviral packaging line for enhanced transfer into primary human T lymphocytes). This vector system in which the cDNA encodes a single-chain antibody (scFv) specific for a tumour-associated antigen can be used for gene therapy of cancer, especially by isolating T cells from the patient's blood, stimulating the T cells in vitro, introducing the retroviral vector into the T cells and returning the T cells to the patient, preferably by intravenous injection

XX Sequence 483 AA;

Query Match 68.6%; Score 1675; DB 2; Length 483;  
Best Local Similarity 69.2%; Pred. No. 1.4e-107;  
Matches 334; Conservative 41; Mismatches 86; Indels 22; Gaps 6;

QY 1 MAMWTLFLMAAK-VPKQIQVQSGPELKKPGEYTKISCKASGYFTYGMWVQAQAP 59  
DB 1 MDWIMRILFLVAGATGASHSEVQLOESGALVRRPSSVKISCKASGYAFSSYGMWVQAQRP 60  
QY 60 GCGKMMGWINTSGESTFADDFKGRFDSLETSANTAYQINNLKSEDMATYFCARMEY 119  
DB 61 GCGLEWIGQIWPQGDPTNNGKFKGKATLTADSSSTAYVQLSLSEDSAAVYFCARRET 120  
QY 120 ----YHGVYWGQGTIVTVSSGGSGS--GGSGSGSGSDIQLTQSHKFLSTSVGDRVS 172  
DB 121 TTVGPTTYANDYWGQGTIVTVSSGGSGS--GGSGSGSGSDIQLTQSHKFLSTSVGDRVS 180  
QY 173 ITCASQDV-YNA---VAMVQKQKPGSKKLITYSASSRYTGVPSRFSGSGGPFTFTIS 228  
DB 181 ISCASQSDVDYDGSYLNWYQQLFGQPPKLLIDASNLVSGIPRFSGSGSGDFTLNIN 240  
QY 229 SVQAEALAVYFCQGHFRPTFGSGTKLEIKALEISNSVMYFSSVVPVLQ 278  
DB 241 PVEKVDATATYHCQOSTEDPWTFGGKTLEIKALEISNSVMYFSSVVPVLQ 300  
QY 279 KVNSTTKPVLRTSPVHPGTSPQRPEDCRPGSVKGTGLDFLEDPKCYLLDGLIFT 338  
DB 301 KVNSTTKPVLRTSPVHPGTSPQRPEDCRPGSVKGTGLDFLEDPKCYLLDGLIFT 360  
QY 339 YGVITLALYRAKFSRSAAETPAANLQDPNQLYNELNGRREYDYLEKKRADPDMGKQ 398  
DB 361 YGVITLALYRAKFSRSAAETPAANLQDPNQLYNELNGRREYDYLEKKRADPDMGKQ 420  
QY 399 RRRNPOGVNNAQKDMAEAYSEITGKERRRGGKHDGLYQGLSTATKDTYDALHMOTL 458  
DB 421 RRRNPOGVNNAQKDMAEAYSEITGKERRRGGKHDGLYQGLSTATKDTYDALHMOTL 480  
QY 459 APR 461  
DB 481 APR 483

RESULT 3

Query Match	62.54	Score 1526	DB 2	Length 443
Best Local Similarity	65.54	Pred. No. 2.6e-97		
Matches 305	Conservative 45	Mismatches 88	Indels 28	Gaps 7
Qy	1	MAWVTLFLPMAAK-VPKQIQIVQSGPELKKPGETVYKISCKASGYPFTYGNMNVKQAP	59	
Dd	1	MGNSCIHLFLAATATATGVHSQVQLQDSGAEILVKFGQASVKLSCKASGYFTFSYMMHNVKQRP	60	
Qy	60	GQGLKMGWINTSTGESLPADDFKGRDFEFLFSANTAYIQINNLKSEDMATYFCARMEV	119	
Dd	61	GRGLEWIGRIDPNSGGTKYNEKFKSKATLLVDPKRSSTAYMQLSLTSDESAVYYCARYDY	120	
Qy	120	Y-HGVYVWQGGTITVYSSGGSGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSYTCAS	178	
Dd	121	YGSSTYDPYQGGTITVYSSGGSGSGGGSGGGSQAVGIC-ESALITISPGETVTLTCRSS	179	
Qy	179	QDYY---NAVAWYQOKPKGQSPKLLIYASASRYTVGSPSFRTSGSGGPDFTTISVQAEDL	235	
Dd	180	TGAVTTSNVAWYQEKRDHLFTGLIGSTNNRAGVPAFPFGSLIGDKAALLTGAQREDE	239	
Qy	236	AVYFCQGHFTPTPTFGSGTKLEIKALEISNSVMYFSSVPEVLQKNSTTTKPVLRTSPV	295	
Dd	240	AIYFCALMYNSMHWFFGGTITLV--LE-----BETTKPVLRTSPV	279	
Qy	296	HPRTSGQPKRPEQCRPGSVKYGGLDPLDLPKLCYLLDGLFTYGVITLALYRAKFSR	355	
Dd	280	HPRTSGQPKRPEQCRPGSVKYGGLD--DPKLCYLLDGLFTYGVITLALYRAKFSR	337	
Qy	356	AETPAANLQDPNQLYNEILNLRREYDVLEKKRAPDPMGKQQRRRNPQEGVYNALQDK	415	
Dd	338	AETPAANLQDPNQLYNEILNLRREYDVLEKKRAPDPMGKQQRRRNPQEGITVYALQDN	397	
Qy	416	MAEYSEIGTKGRRRGKGDGLYQGLSTATKDTYDALHMQTLAPR	461	
Dd	398	MAEYSEIGTKGRRRGKGDGLYQGLSTATKDTYDALHMQTLAPR	443	
RESULT 4				
AAW24027				
ID	AAW24027	standard; protein; 428 AA.		
AC	AAW24027;			
XX	XX			
DT	25-MAR-2003	(revised)		
DT	04-MAR-1998	(first entry)		
XX	XX			
DE	Single chain antigen hybrid receptor partial sequence.			
XX	XX			
KW	Hybrid receptor; single chain antigen; gene therapy; diagnosis;			
KW	signal conduction; receptor; control region.			
OS	Synthetic.			
OS	Homo sapiens.			
XX	XX			
Key	Location/Qualifiers			
FT	1..4			
FT	/note="partial leader sequence"			
FT	5..428			
FT	/note="single chain antigen hybrid receptor"			
FT	Region			
FT	/label="VH_B_1-8			
FT	/note="variable heavy chain region of B1-8 antibody "			
FT	125..139			
FT	/label="(Gly_Ser)3			
FT	/note="linker region"			
FT	140..249			
FT	/label="VL_B_1-8			
FT	/note="variable light chain region of B1-8 antibody"			
FT	250..294			
FT	/label="CD_8alpha_hinge			
FT	295..315			
FT	/label="transmembrane_domain			
FT	311..428			







```

XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= ompa_signal_peptide
FT Peptide 22..29
FT /label= FLAG_peptide_and_enterokinase_cleavage_site
FT Domain 33..151
FT /label= FRP5_heavy_chain_variable_domain
FT Peptide 152..166
FT /label= Linker
FT Domain 167..274
FT /label= FRP5_light_chain_variable_domain
FT Protein 276..397
FT /label= ETA_252-613
XX EP502812-A1.
XX PD 09-SEP-1992.
XX PF 27-JAN-1992; 92EP-00810056.
XX PR 05-FEB-1991; 91EP-00810079.
XX (CIBA ) CIBA GEIGY AG.
XX PI Wels WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX N-PSDB; AAQ28262.
XX DR WPI; 1992-302096/37.
XX PT Recombinant antibodies directed to growth factor receptor C-ERBB-2 - for
XX PT diagnosing and treating tumours expressing C-ERBB-2 e.g. breast or
XX PT ovarian tumours.
XX PS Disclosure: Page 47-52; 67pp; English.
XX CC The sequences given in AAR26982-3 contain part of the exotoxin A (ETA)
XX CC sequence corresponding to positions 252-613 of the full exotoxin A
XX CC sequence. These sequences are encoded by Fv(FRP5)-ETA fusion genes. The
XX CC ETA sequence was used as a marker gene so that E. coli transformed with
XX CC the fusion gene could be identified. The fusion genes were expressed in
XX CC E. coli and the antibodies were extracted. These recombinant antibodies
XX CC can be used for the qualitative and quantitative determination of tumours
XX CC 2. This can be used for monitoring or in-vivo localisation of tumours
XX CC overexpressing c-erbB-2. (Updated on 25-MAR-2003 to correct PN field.)
XX CC (Updated on 24-OCT-2003 to standardise OS field)
XX SQ Sequence 637 AA;

```

```

Query Match 53.4%; Score 1304; DB 2; Length 637;
Best Local Similarity 97.6%; Pred. No. 9e-82;
Matches 242; Conservative 3; Mismatches 3; Indels 0; Gaps 0;
QY 15 KVPKQIQVOSGPELKKPEETVKISCKASGYPTNYGNMVKQAPGGGLKMMGWINTSGSTF 74
DB 29 KLASGVQIQOSGPELKKPEETVKISCKASGYPTNYGNMVKQAPGGGLKMMGWINTSGSTF 88
QY 75 ESTRADPKRFPDSLETSANTAVYQINNLSKSEDMATYFCARWEVYHGYPYWGQTTVT 134
DB 89 ESTRADPKRFPDSLETSANTAVYQINNLSKSEDMATYFCARWEVYHGYPYWGQTTVT 148
QY 135 VSSGGGGGGGGGGSDIQLTQSHKPLSTSVGDRVSTTCKASQDYNNAVNYOQKPGQ 194
DB 149 VSSGGGGGGGGGGSDIQLTQSHKPLSTSVGDRVSTTCKASQDYNNAVNYOQKPGQ 208
QY 195 SPKLLIYSASSRYTGVSRFTYSGSGPDTFTTISVQAEADLAYVFCQOHRTPTTSGSGT 254
DB 209 SPKLLIYSASSRYTGVSRFTYSGSGPDTFTTISVQAEADLAYVFCQOHRTPTTSGSGT 268
QY 255 KLEIKALE 262
DB 269 KLEIKALE 276

```

```

RESULT 8
AAR26980
ID AAR26980 standard; protein; 711 AA.
XX
XX AAR26980;
AC 25-MAR-2003 (revised)
XX 11-FEB-1993 (first entry)
XX
XX Fv(FRP5)-phoA recombinant antibody.
XX Monoclonal antibody; light chain; heavy chain; tumour; phoA;
XX variable region; alkaline phosphatase; c-erbB-2; ss.
XX Synthetic.
XX
XX Key Location/Qualifiers
FH Peptide 1..21
FT /label= ompa_signal_peptide
FT Domain 23..141
FT /label= FRP5_heavy_chain_variable_domain
FT Peptide 142..156
FT /label= Linker
FT Domain 157..264
FT /label= FRP5_light_chain_variable_domain
FT Protein 265..711
FT /label= phoA
XX
XX EP502812-A1.
XX PD 09-SEP-1992.
XX PF 27-JAN-1992; 92EP-00810056.
XX PR 05-FEB-1991; 91EP-00810079.
XX (CIBA ) CIBA GEIGY AG.
XX PI Wels WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX N-PSDB; AAQ28257.
XX DR WPI; 1992-302096/37.
XX PT Recombinant antibodies directed to growth factor receptor C-ERBB-2 - for
XX PT diagnosing and treating tumours expressing C-ERBB-2 e.g. breast or
XX PT ovarian tumours.
XX PS Disclosure: Page 34-40; 67pp; English.
XX CC The sequence given is the single chain recombinant antibody designated
XX CC Fv(FRP5)-phoA. The alkaline phosphatase gene (phoA) was used as a marker
XX CC gene so that E. coli transformed with the fusion gene could be
XX CC identified. The fusion gene was expressed in E. coli and the antibody was
XX CC extracted. This recombinant antibody can be used for the qualitative and
XX CC quantitative determination of c-erbB-2. This can be used for monitoring
XX CC or in-vivo localisation of tumours overexpressing c-erbB-2. (Updated on
XX CC 25-MAR-2003 to correct PN field.)
XX SQ Sequence 711 AA;

```

```

Query Match 53.4%; Score 1303; DB 2; Length 711;
Best Local Similarity 94.2%; Pred. No. 1.2e-81;
Matches 245; Conservative 2; Mismatches 1; Indels 12; Gaps 1;
QY 19 QIQVOSGPELKKPEETVKISCKASGYPTNYGNMVKQAPGGGLKMMGWINTSGESTF 78
DB 23 QIQVOSGPELKKPEETVKISCKASGYPTNYGNMVKQAPGGGLKMMGWINTSGESTF 82
QY 79 ADDFKGRFPDSLETSANTAVYQINNLSKSEDMATYFCARWEVYHGYPYWGQTTVTSSG 138
DB 83 ADDFKGRFPDSLETSANTAVYQINNLSKSEDMATYFCARWEVYHGYPYWGQTTVTSSG 142

```

```

QY 139 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAVMYOQKPGQSPRL 198
DB 143 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAVMYOQKPGQSPRL 202
QY 199 LIYSASSRYTGVPSRFPTSGSGDPFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTKLEI 258
DB 203 LIYSASSRYTGVPSRFPTSGSGDPFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTKLEI 262
QY 259 KALEISNSVMYFSSVVPVUQ 278
DB 263 KALE-----PVLE 270

RESULT 9
AAR95053 standard; protein; 530 AA.
ID AAR95053
AC AAR95053;
XX
XX 16-OCT-2003 (revised)
DT 18-AUG-1996 (first entry)
DE scFv (FRP5)-DETA-DGAL4 multidomain protein.
XX scFv (FRP5)-DETA-DGAL4 multidomain protein.
XX Nucleic acid transfer system; gene transfer; gene therapy;
KW cell targeting; multidomain protein; vector; cancer; exotoxin A; DETA;
KM single chain antibody; scFv; GAL4.
XX
XX Mus; sp.
OS Pseudomonas; aeruginosa.
OS Saccharomyces cerevisiae.
OS Chimeric.
OS Synthetic.
XX
XX Location/Qualifiers
FH 1.8
FT /label= FLAG_epitope
FT 9..17
FT /label= Spacer
FT 18..257
FT /label= ScFv (FRP5)
FT 258..260
FT /label= Spacer
FT 261..375
FT /label= ETR
FT /note= "amino acids 252-366 of exotoxin-A"
FT Peptide
FT 376
FT /label= Spacer
FT 377..522
FT /label= GAL4
FT /note= "amino acids 2-147 of yeast GAL4"
FT Peptide
FT 523..530
FT /label= Spacer
FT /note= "endoplasmic reticulum retention peptide"
XX
XX WO9613599-A1.
XX
XX 09-MAY-1996.
XX
XX 31-OCT-1995; 95WO-EP004270.
XX
XX 01-NOV-1994; 94BP-00810627.
XX
XX (WELIS/) WELIS W.
XX
XX Wels W, Fominaya J;
XX
XX MPI; 1996-239505/24.
XX
XX N-PsDB; AAT29409.
XX
XX Nucleic acid transfer system for gene therapy, e.g. against cancer -
XX includes toxin translocation domain to target nucleic acid to specific
XX cell.

```

```

XX
PS Claim 7; Page 59-61; 106pp; English.
XX
CC A multidomain protein (AAR95053) has a FLAG epitope, a single chain
CC antibody, scFv, of monoclonal antibody FRP5 (raised against human tumour
CC cell HBR2 antigen) that acts as a ligand domain, a non-cytotoxic portion
CC of Pseudomonas aeruginosa exotoxin A acting as a translocation domain and
CC the DNA binding domain of yeast GAL4. It is the product of a fusion gene
CC (AAT29409) and can be expressed in E. coli (resulting in removal of ompA
CC signal peptide). It is used with an effector nucleic acid that comprises
CC e.g. a gene to be delivered to a cell and a cognate structure for the
CC GAL4 DNA binding domain. This provides a novel means of nucleic acid
CC transfer, suitable for gene therapy. (Updated on 16-OCT-2003 to
CC standardise OS field)
XX
SQ Sequence 530 AA;
XX
XX Query Match 53.3%; Score 1302; DB 2; Length 530;
XX Best Local Similarity 99.2%; Pred. No. 1e-81;
XX Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
QY 19 QIQLVQSGPELKRPETVKISCKASGYPTVYGMWVQAQGLKMKMGWINTSGESTF 78
DB 18 QVQLQQSGPELKRPETVKISCKASGYPTVYGMWVQAQGLKMKMGWINTSGESTF 77
QY 79 ADDPKGRDFSLFETSNANTAYVQINNLIKSEDMATYFCARMEYHGVVPYMGQTTVTVSSG 138
DB 78 ADDPKGRDFSLFETSNANTAYVQINNLIKSEDMATYFCARMEYHGVVPYMGQTTVTVSSG 137
QY 139 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAVMYOQKPGQSPRL 198
DB 138 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAVMYOQKPGQSPRL 197
QY 199 LIYSASSRYTGVPSRFPTSGSGDPFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTKLEI 258
DB 198 LIYSASSRYTGVPSRFPTSGSGDPFTFTISSVOAEDLAVYFCQOHFRTPFTFGSGTKLEI 257
QY 259 KALE 262
DB 258 KALE 261

RESULT 10
AAM05136 standard; protein; 651 AA.
ID AAM05136
AC AAM05136;
XX
XX 29-JAN-1997 (first entry)
DT
XX
XX scFv (FRP5)-ETRA fusion protein.
DE scFv (FRP5)-ETRA fusion protein.
XX
XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pSW202-5; cancer; therapy;
KM antitumour; exotoxin A; ETR.
XX
XX Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
XX Key
FH 1..21
FT /label= Sig_peptide
FT /note= "ompA signal peptide"
FT 22..38
FT /label= Spacer
FT 39..278
FT /label= scFv (FRP5)
FT 279..289
FT /label= Spacer
FT 290..651
FT /label= ETR

```



QY 139 GGGSGGGGGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAVYQKPGQSPKL 198  
 DB 159 GGGSGGGGGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAVYQKPGQSPKL 218  
 QY 199 LIYSASSRYTGVPSRFTTSGSGSPDFTFTTSSVQAEIDLAVYFCQOHFRPTFTGSGTKLEI 258  
 DB 219 LIYSASSRYTGVPSRFTTSGSGSPDFTFTTSSVQAEIDLAVYFCQOHFRPTFTGSGTKLEI 278  
 QY 259 KALE 262  
 DB 279 KALE 282

## RESULT 12

AAW05140  
 ID AAW05140 standard; protein; 892 AA.  
 AC AAW05140;  
 XX 29-JAN-1997 (first entry)  
 DE scFv2(225/FRP5)-ETA.  
 XX  
 KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;  
 KW epidermal growth factor; receptor; plasmid pMS338-225-5; cancer;  
 KW exotoxin A; ETA; antitumour.  
 XX  
 OS Mus; BP.  
 OS Pseudomonas; aeruginosa.  
 OS Synthetic.  
 OS Chimeric.  
 XX  
 FH Key  
 FT Peptide  
 FT 1..21 Location/Qualifiers  
 FT /label= Sig\_peptide  
 FT /note= "ompa signal peptide"  
 FT Peptide  
 FT 22..38  
 FT /label= Spacer  
 FT Region  
 FT 39..278  
 FT /label= scFv(225)  
 FT Peptide  
 FT 279..289  
 FT /label= Spacer  
 FT Region  
 FT 290..404  
 FT /label= ETA  
 FT /note= "exotoxin A amino acids 252-366"  
 FT Peptide  
 FT 405..407  
 FT /label= Spacer  
 FT Region  
 FT 408..647  
 FT /label= scFv(FRP5)  
 FT Peptide  
 FT 648..658  
 FT /label= Spacer  
 FT Region  
 FT 659..892  
 FT /label= ETA  
 FT /note= "endotoxin-A amino acids 380-613"  
 XX  
 PN EP739984-A1.  
 XX 30-OCT-1996.  
 PD 26-APR-1995;  
 PE 95EP-00106275.  
 XX 26-APR-1995;  
 PR 95EP-00106275.  
 XX 26-APR-1995;  
 PA (SANT-) SAN TUMORFORSCHUNGS GMBH.  
 PI (SANT-) SAN TUMORFORSCHUNGS GMBH.  
 PI Weis W, Schmidt M, Groner B;  
 XX MPI; 1996-478748/48.  
 DR N-P8DB; AAT42040.  
 XX Bivalent fusion proteins that bind epidermal growth factor receptor or  
 PT analogues - and comprise at least two different cell surface binding  
 PT domain(s), useful for tumour therapy.

XX Example 11; Page 31-33; 52pp; English.  
 PS  
 XX scFv2(225/FRP5)-ETA (AAW05140) comprises the single-chain binding region  
 CC of murine monoclonal antibody 225 (specific for human epidermal growth  
 CC factor receptor, see also AAW05133) joined to portions of exotoxin A from  
 CC Pseudomonas aeruginosa and to the single-chain binding region of murine  
 CC monoclonal antibody FRP5 (specific for human epidermal growth factor  
 CC receptor erbB-2, see also AAW05134). It is encoded by plasmid pMS238-225-  
 CC 5 (AAT42040). This plasmid can be utilized in the prodn. of the bivalent  
 CC fusion protein in bacterial (esp. E. coli) host cells. Such fusion  
 CC proteins (see also AAW05138-44) are useful as antitumour agents  
 XX  
 SQ Sequence 892 AA;

Query Match 53.3%; Score 1302; DB 2; Length 892;  
 Best Local Similarity 99.2%; Pred. No. 1,8e-81;  
 Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 19 QIQLVQSGPELKKPGETVYKISCKASGYPTNYGMNMYVQAAPGQGLKMMGWINTSGESTF 78  
 DB 408 QVQLQQSGPELKKPGETVYKISCKASGYPTNYGMNMYVQAAPGQGLKMMGWINTSGESTF 467  
 QY 79 ADFFKGRFDSLETSANTAYVQINNLSKEDMATYFCARWEYVHGVPYWGQGTIVTVSSG 138  
 DB 468 ADFFKGRFDSLETSANTAYVQINNLSKEDMATYFCARWEYVHGVPYWGQGTIVTVSSG 527  
 QY 139 GGGSGGGGGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAVYQKPGQSPKL 198  
 DB 528 GGGSGGGGGGGSDIQLTOSHKFLSTSVGDRVSIITCKASQDVYNAVAVYQKPGQSPKL 587  
 QY 199 LIYSASSRYTGVPSRFTTSGSGSPDFTFTTSSVQAEIDLAVYFCQOHFRPTFTGSGTKLEI 258  
 DB 588 LIYSASSRYTGVPSRFTTSGSGSPDFTFTTSSVQAEIDLAVYFCQOHFRPTFTGSGTKLEI 647  
 QY 259 KALE 262  
 DB 648 KALE 651

## RESULT 13

AAW05143  
 ID AAW05143 standard; protein; 892 AA.  
 AC AAW05143;  
 XX 29-JAN-1997 (first entry)  
 DE scFv2(FRP5/FRP5)-ETA (version 2).  
 XX  
 KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;  
 KW epidermal growth factor; receptor; plasmid pMS338-5-5; cancer;  
 KW exotoxin A; ETA; antitumour.  
 XX  
 OS Mus; BP.  
 OS Pseudomonas; aeruginosa.  
 OS Synthetic.  
 OS Chimeric.  
 XX  
 FH Key  
 FT Peptide  
 FT 1..21 Location/Qualifiers  
 FT /label= Sig\_peptide  
 FT /note= "ompa signal peptide"  
 FT Peptide  
 FT 22..38  
 FT /label= Spacer  
 FT Region  
 FT 39..278  
 FT /label= scFv(FRP5)  
 FT Peptide  
 FT 279..289  
 FT /label= Spacer  
 FT Region  
 FT 290..404  
 FT /label= ETA  
 FT /note= "exotoxin A amino acids 252-366"  
 FT Peptide  
 FT 405..407

```

FT      /label= Spacer
FT      408..647
FT      /label= scFv(FRP5)
FT      Peptide
FT      648..658
FT      /label= Spacer
FT      Region
FT      659..892
FT      /label= ETA
FT      /note= "exotoxin A amino acids 380-613"
XX      EP739984-A1.
XX      30-OCT-1996.
XX      26-APR-1995; 95EP-00106275.
XX      26-APR-1995; 95EP-00106275.
XX      (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX      Wels W, Schmidt M, Groner B;
XX      WPI; 1996-478748/48.
XX      N-PSDB; AAT42043.
XX      Bivalent fusion proteins that bind epidermal growth factor receptor or
XX      PT analogues - and comprise at least two different cell surface binding
XX      PT domain(s), useful for tumour therapy.
XX      Example 12; Page 40-42; 52pp; English.
XX      scFv2(FRP5/FRP5)-ETA version 2 (AAW05139) includes 2 copies of the single
XX      CC -chain binding region of murine monoclonal antibody FRP5 (specific for
XX      CC human epidermal growth factor receptor erbB-2, see also AAW05134) joined
XX      CC to portions of exotoxin A from Pseudomonas aeruginosa. It is encoded by
XX      CC plasmid pMS238-5-5 (AAT42043). This plasmid can be utilised in the prodn.
XX      CC of the bivalent fusion protein in bacterial (esp. E. coli) host cells.
XX      CC Such fusion proteins (see also AAW05138-44) are useful as antitumour
XX      CC agents
XX      SQ Sequence 892 AA;
XX
Query Match      53.3%; Score 1302; DB 2; Length 892;
Best Local Similarity 99.2%; Pred. No. 1,8e-81;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
XX
QY      19 QIQLVQSGPELKKPGETVYKISCKASGYPTNYGMNWKQAPGQGLKMMGWINTSTGESTF 78
DB      :|||||
XX      39 QVQLQQSGPELKKPGETVYKISCKASGYPTNYGMNWKQAPGQGLKMMGWINTSTGESTF 98
QY      79 ADDPKRFPDFSLSTANTAYLIQINNLKSEDMATYFCARWEYTHGVVPRWGQGTIVTVSSG 138
DB      :|||||
XX      99 ADDPKRFPDFSLSTANTAYLIQINNLKSEDMATYFCARWEYTHGVVPRWGQGTIVTVSSG 158
QY      139 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDYNNAVAWYQKRGQSPKL 198
DB      :|||||
XX      159 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDYNNAVAWYQKRGQSPKL 218
QY      199 LIYSASSRYTVGPSRFTSSGGSPDPTFTISSVOAEDLAIVFCQGHFPTFTFGSGTKLEI 258
DB      :|||||
XX      219 LIYSASSRYTVGPSRFTSSGGSPDPTFTISSVOAEDLAIVFCQGHFPTFTFGSGTKLEI 278
QY      259 KALE 262
DB      :|||||
XX      279 KALE 282
XX
RESULT 14
AAW05139
ID      AAW05139 standard; protein; 892. AA.
XX      AC      AAW05139;
XX      DT      29-JAN-1997 (first entry)

```

```

XX      XX      scFv2(FRP5/225)-ETA (version 1).
DE      DE      Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
XX      XX      epidermal growth factor; receptor; plasmid pMS238-5-225; cancer;
XX      XX      exotoxin A; ETA; antitumour.
XX      OS      Mus; sp.
XX      OS      Pseudomonas; aeruginosa.
XX      OS      Synthetic.
XX      OS      Chimeric.
XX      FH      Key
XX      FT      Peptide
XX      FT      Location/Qualifiers
XX      FT      1..21
XX      FT      /label= Sig peptide
XX      FT      /note= "ompA signal peptide"
XX      FT      Peptide
XX      FT      22..38
XX      FT      /label= Spacer
XX      FT      39..278
XX      FT      /label= scFv(FRP5)
XX      FT      Peptide
XX      FT      279..289
XX      FT      /label= Spacer
XX      FT      Region
XX      FT      290..404
XX      FT      /label= ETA
XX      FT      /note= "exotoxin A amino acids 252-366"
XX      FT      405..407
XX      FT      /label= Spacer
XX      FT      Region
XX      FT      408..647
XX      FT      /label= scFv(225)
XX      FT      Peptide
XX      FT      648..658
XX      FT      /label= Spacer
XX      FT      Region
XX      FT      659..892
XX      FT      /label= ETA
XX      FT      /note= "endotoxin-A amino acids 380-613"
XX      EP739984-A1.
XX      30-OCT-1996.
XX      26-APR-1995; 95EP-00106275.
XX      26-APR-1995; 95EP-00106275.
XX      26-APR-1995; 95EP-00106275.
XX      (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX      Wels W, Schmidt M, Groner B;
XX      WPI; 1996-478748/48.
XX      N-PSDB; AAT42039.
XX      Bivalent fusion proteins that bind epidermal growth factor receptor or
XX      PT analogues - and comprise at least two different cell surface binding
XX      PT domain(s), useful for tumour therapy.
XX      Example 11; Page 28-30; 52pp; English.
XX      scFv2(FRP5/225)-ETA (AAW05139) comprises the single-chain binding region
XX      CC of murine monoclonal antibody FRP5 (specific for human epidermal growth
XX      CC factor receptor erbB-2, see also AAW05134) joined to portions of exotoxin
XX      CC A from Pseudomonas aeruginosa and to the single-chain binding region of
XX      CC murine monoclonal antibody 225 (specific for human epidermal growth
XX      CC factor receptor, see also AAW05133). It is encoded by plasmid pMS238-5-
XX      CC 225 (AAT42039). This plasmid can be utilised in the prodn. of the
XX      CC bivalent fusion protein in bacterial (esp. E. coli) host cells. Such
XX      CC fusion proteins (see also AAW05138-44) are useful as antitumour agents
XX      CC
XX      SQ Sequence 892 AA;
XX
Query Match      53.3%; Score 1302; DB 2; Length 892;
Best Local Similarity 99.2%; Pred. No. 1,8e-81;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY      19 QIQLVQSGPELKKPGETVYKISCKASGYPTNYGMNWKQAPGQGLKMMGWINTSTGESTF 78

```

```

DB 39 QVQLQSGPELKKPGETYKISCKASGYPTNYGMNHWQAQPGQGLKMMGMINTSTGSESTF 98
QY 79 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARMEYHGVVPYWGQGTITVSSG 138
DB 99 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARMEYHGVVPYWGQGTITVSSG 158
QY 139 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 198
DB 159 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 218
QY 199 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVQAEIDLAVYFCQGHFRTPFTFGSGTKLEI 258
DB 219 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVQAEIDLAVYFCQGHFRTPFTFGSGTKLEI 278
QY 259 KALE 262
DB 279 KALE 282

```

## RESULT 15

```

AAW05142 ID AAW05142 standard; protein; 895 AA.
XX AC AAW05142;
XX DT 29-JAN-1997 (first entry)
DE scFv2(FRP5/FRP5)-ETA (version 1).
XX KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KM epidermal growth factor; receptor; plasmid pMS242-5-5; cancer;
XX exotoxin A; ETA; antitumour.
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX FH Key
FH FT Location/Qualifiers
FT FT 1..21
FT FT /label= sig_peptide
FT FT /note= "ompA signal peptide"
FT FT 22..44
FT FT /label= Spacer
FT FT 45..159
FT FT /label= ETA
FT FT /note= "exotoxin A amino acids 252-366"
FT FT 160..162
FT FT /label= Spacer
FT FT 163..402
FT FT /label= scFv(FRP5)
FT FT 403..413
FT FT /label= Spacer
FT FT 414..644
FT FT /label= ETA
FT FT /note= "exotoxin A amino acids 380-610"
FT FT 645..646
FT FT /label= Spacer
FT FT 647..886
FT FT /label= scFv(FRP5)
FT FT 887..895
FT FT /label= Spacer
XX PN EP739984-A1.
XX PD 30-OCT-1996.
XX PF 26-APR-1995; 95SEP-00106275.
XX PR 26-APR-1995; 95SEP-00106275.
XX PA (SANT-) SAN TUMORFORSCHUNGS GMBH.

```

```

XX PI Wele W, Schmidt M, Groner B;
XX DR WPI: 1996-476746/48.
XX DR N-PSDB: AAT42042.
XX PT Bivalent fusion proteins that bind epidermal growth factor receptor or
XX PT analogues - and comprise at least two different cell surface binding
XX PT domain(s), useful for tumour therapy.
XX PS Example 12; Page 37-39; 52pp; English.
XX CC scFv2(FRP5/FRP5)-ETA version 1 (AAW05142) includes 2 copies of the single
XX CC -chain binding region of murine monoclonal antibody FRP5 (specific for
XX CC human epidermal growth factor receptor erbB-2, see also AAW05134) joined
XX CC to portions of exotoxin A from Pseudomonas aeruginosa. It is encoded by
XX CC plasmid pMS242-5-5 (AAT42042). This plasmid can be utilised in the prodn.
XX CC of the bivalent fusion protein in bacterial (esp. E. coli) host cells.
XX CC Such fusion proteins (see also AAW05138-44) are useful as antitumour
XX CC agents
XX SQ Sequence 895 AA;

```

Query Match 53.3%; Score 1302; DB 2; Length 895;  
 Best Local Similarity 99.2%; Pred. No. 1.8e-81;  
 Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

QY 19 QIQLVQSGPELKKPGETYKISCKASGYPTNYGMNHWQAQPGQGLKMMGMINTSTGSESTF 78
DB 163 QVQLQSGPELKKPGETYKISCKASGYPTNYGMNHWQAQPGQGLKMMGMINTSTGSESTF 222
QY 79 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARMEYHGVVPYWGQGTITVSSG 138
DB 223 ADDPKGRDPSLETSANTAYLQINNLSKEDMATYFCARMEYHGVVPYWGQGTITVSSG 282
QY 139 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 198
DB 283 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQKPGQSPKL 342
QY 199 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVQAEIDLAVYFCQGHFRTPFTFGSGTKLEI 258
DB 343 LIYSASSRYTGVPSRFTSGSGPDFTFTISSVQAEIDLAVYFCQGHFRTPFTFGSGTKLEI 402
QY 259 KALE 262
DB 403 KALE 406

```

Search completed: April 20, 2005, 07:04:52  
 Job time : 70.8241 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 07:01:08 ; Search time 20.9082 Seconds  
(without alignments)  
1645.921 Million cell updates/sec

Title: US-09-596-774-6  
Perfect score: 2441  
Sequence: 1 MAWWTLLFLMAAKVPKQI.....ISTATKDTYDALHMQLTAPR 461

Scoring table: BLASTSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1308	53.6	615 4 US-08-840-713-35	Sequence 35, Appl
2	1308	53.6	617 4 US-08-840-713-37	Sequence 37, Appl
3	1307	53.5	637 1 US-08-235-838-14	Sequence 14, Appl
4	1307	53.5	637 2 US-08-465-473B-14	Sequence 14, Appl
5	1303	53.4	711 1 US-08-235-838-7	Sequence 7, Appl
6	1303	53.4	711 2 US-08-465-473B-7	Sequence 7, Appl
7	1302	53.3	530 4 US-08-840-713-2	Sequence 2, Appl
8	1284	52.6	241 1 US-08-235-838-5	Sequence 5, Appl
9	1284	52.6	241 2 US-08-465-473B-5	Sequence 5, Appl
10	981.5	40.2	250 1 US-08-133-804-2	Sequence 8, Appl
11	981.5	40.2	250 1 US-08-461-184-8	Sequence 8, Appl
12	981.5	40.2	250 1 US-08-463-675-8	Sequence 8, Appl
13	981.5	40.2	250 1 US-08-464-589-8	Sequence 8, Appl
14	981.5	40.2	250 1 US-08-461-838-2	Sequence 2, Appl
15	981.5	40.2	250 2 US-08-461-386-2	Sequence 2, Appl
16	975.5	40.0	622 2 US-08-356-786-16	Sequence 16, Appl
17	903.5	37.0	240 1 US-08-488-113B-148	Sequence 148, App
18	903.5	37.0	240 1 US-08-477-484B-148	Sequence 148, App
19	903.5	37.0	240 2 US-08-466-360-148	Sequence 148, App
20	903.5	37.0	240 3 US-08-839-765-148	Sequence 148, App
21	903.5	37.0	240 3 US-09-136-389-148	Sequence 148, App
22	903.5	37.0	240 3 US-09-610-838-148	Sequence 148, App
23	903.5	37.0	240 4 US-09-711-485-148	Sequence 148, App
24	901.5	36.9	267 3 US-09-485-737B-2	Sequence 2, Appl
25	901.5	36.9	267 4 US-10-071-485-2	Sequence 2, Appl
26	896	36.7	365 3 US-08-875-811-53	Sequence 53, Appl
27	891	36.5	366 3 US-08-875-811-55	Sequence 55, Appl

28	889.5	36.4	541 3 US-09-485-737B-85	Sequence 85, Appl
29	889.5	36.4	541 4 US-10-071-485-85	Sequence 85, Appl
30	888.5	36.4	711 3 US-09-485-737B-90	Sequence 90, Appl
31	888.5	36.4	711 4 US-10-071-485-90	Sequence 90, Appl
32	888	36.4	284 3 US-09-184-658-40	Sequence 40, Appl
33	888	36.4	284 4 US-09-504-262D-40	Sequence 40, Appl
34	885.5	36.3	259 4 US-09-419-788-115	Sequence 115, App
35	878.5	36.0	534 2 US-08-356-786-10	Sequence 10, Appl
36	851	34.9	240 4 US-10-092-246-35	Sequence 35, Appl
37	851	34.9	240 4 US-10-092-246-35	Sequence 35, Appl
38	851	34.9	240 4 US-10-096-246A-35	Sequence 35, Appl
39	851	34.9	240 4 US-10-096-246A-37	Sequence 37, Appl
40	848	34.7	240 3 US-09-485-737B-91	Sequence 91, Appl
41	848	34.7	240 4 US-10-071-485-91	Sequence 91, Appl
42	843	34.5	240 4 US-10-096-246A-36	Sequence 36, Appl
43	841.5	34.5	249 2 US-08-797-689-18	Sequence 18, Appl
44	841.5	34.5	249 4 US-09-984-186-18	Sequence 18, Appl
45	839	34.4	240 4 US-10-092-246-37	Sequence 37, Appl

# ALIGNMENTS

```

RESULT 1
US-08-840-713-35
; Sequence 35, Application US/08840713
; Patent No. 6498233
;
GENERAL INFORMATION:
; APPLICANT: WELS, Winfried, Dr.
; APPLICANT: FOYMINAYA, Jesus
; TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESS: Nikolaus, Marmelstein, Murray & Oram LLP
; STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-5701
;
COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840, 713
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kites, Monica Chin
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: 1614-7014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638 - 5000
; TELEFAX: (202) 638 - 4810
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-840-713-35
Query Match 53.6%; Score 1308; DB 4; Length 615;
Best Local Similarity 96.5%; Pred. No. 2.5e-100;
Matches 245; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
QY 19 QIOLVSGPELKKKGGTYKISCKASGYPTNYGNWVKQAPGGLKMMGWINTSGESTF 78
DB 362 QVQLQSGPELKKKGGTYKISCKASGYPTNYGNWVKQAPGGLKMMGWINTSGESTF 421
QY 79 ADKFGKDFSLSTANTAYLQINNLSKSDMATYFCARWEVYHGVVPYGGCTITVYSSG 138

```

Db 422 ADPFKGRFDSLETSANTAYLQINNLKSEDMATYFCARWEYHGVYPYWGQTTVTSSG 481  
QY 139 GGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDYNNAVAWYQKPGQSPRL 198  
Db 482 GGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDYNNAVAWYQKPGQSPRL 541  
QY 199 LIYSASRYTGVPSRFTGSGGPDFTTISVQAEIDLAVYFCQHFRTPTFTGSGTKLEI 258  
Db 542 LIYSASRYTGVPSRFTGSGGPDFTTISVQAEIDLAVYFCQHFRTPTFTGSGTKLEI 601  
QY 259 KALEINSVWYFSS 272  
Db 602 KALEIDLSERRRFA 615

## RESULT 2

US-08-840-713-37  
Sequence 37, Application US/08840713  
Patent No. 6498233  
GENERAL INFORMATION:  
APPLICANT: WELLS, Winfried, Dr.  
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Matmellestein, Murray & Oram LLP  
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/840,713  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Killea, Monica Chih  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: 1614-7014  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638 - 4810  
TELEFAX: (202) 638 - 4810  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 617 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-840-713-37

Query Match 53.6%; Score 1308; DB 4; Length 617;  
Best Local Similarity 96.5%; Pred. No. 2.5e-100;  
Matches 245; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 19 QIOLVQSGPELKKRGEYVTKISCKASGYPTNYGNMVKQAPGQGLKMMGWINTSTESTF 78  
Db 364 QVLOQSGPELKKRGEYVTKISCKASGYPTNYGNMVKQAPGQGLKMMGWINTSTESTF 423  
QY 79 ADPFKGRFDSLETSANTAYLQINNLKSEDMATYFCARWEYHGVYPYWGQTTVTSSG 138  
Db 424 ADPFKGRFDSLETSANTAYLQINNLKSEDMATYFCARWEYHGVYPYWGQTTVTSSG 483  
QY 139 GGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDYNNAVAWYQKPGQSPRL 198  
Db 484 GGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDYNNAVAWYQKPGQSPRL 543  
QY 199 LIYSASRYTGVPSRFTGSGGPDFTTISVQAEIDLAVYFCQHFRTPTFTGSGTKLEI 258

Db 544 LIYSASRYTGVPSRFTGSGGPDFTTISVQAEIDLAVYFCQHFRTPTFTGSGTKLEI 603  
QY 259 KALEINSVWYFSS 272  
Db 604 KALEIDLSERRRFA 617

## RESULT 3

US-08-235-838-14  
Sequence 14, Application US/08235838  
Patent No. 5571894  
GENERAL INFORMATION:  
APPLICANT: WELLS, Winfried S.  
APPLICANT: Hynes, Nancy E.  
APPLICANT: Harwerth, Ina-Maria  
APPLICANT: Groner, Bernd  
APPLICANT: Hardman, No. 5571894man  
APPLICANT: Zwickl, Markus  
TITLE OF INVENTION: Recombinant Antibodies Specific for a  
TITLE OF INVENTION: Growth Factor Receptor  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,838  
FILING DATE: TBA  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/828,832  
FILING DATE: 31-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 91-810079.3  
FILING DATE: 05-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 14:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 637 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-838-14

Query Match 53.5%; Score 1307; DB 1; Length 637;  
Best Local Similarity 98.0%; Pred. No. 3.1e-100;  
Matches 243; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 15 KVPKQIOLVQSGPELKKRGEYVTKISCKASGYPTNYGNMVKQAPGQGLKMMGWINTSTG 74  
Db 29 KLSQVLOQSGPELKKRGEYVTKISCKASGYPTNYGNMVKQAPGQGLKMMGWINTSTG 88  
QY 75 ESTFADDFKGRFDSLETSANTAYLQINNLKSEDMATYFCARWEYHGVYPYWGQTTVT 134  
Db 89 ESTFADDFKGRFDSLETSANTAYLQINNLKSEDMATYFCARWEYHGVYPYWGQTTVT 148  
QY 135 VSSGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDYNNAVAWYQKPGQ 194  
Db 149 VSSGGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDYNNAVAWYQKPGQ 208

QY 195 SPKLIYASASSRYTGVPSRFTGSGSPDFTTISVQAEADLAVVFCQOHFRTPTFGSGT 254  
 DB 209 SKKLIYASASSRYTGVPSRFTGSGSPDFTTISVQAEADLAVVFCQOHFRTPTFGSGT 268  
 QY 255 KLEIKALE 262  
 DB 269 KLEIKALE 276

RESULT 4  
 US-08-465-473B-14  
 ; Sequence 14, Application US/08465473B  
 ; Patent No. 5939531

GENERAL INFORMATION:  
 APPLICANT: Weis, Winfried S.  
 APPLICANT: Hynes, Nancy E.  
 APPLICANT: Harwerth, Ina-Maria  
 APPLICANT: Groner, Bernd  
 APPLICANT: Hardman, No. 5939531man  
 APPLICANT: Zwickl, Markus  
 TITLE OF INVENTION: Recombinant Antibodies Specific for a  
 TITLE OF INVENTION: Growth Factor Receptor  
 NUMBER OF SEQUENCES: 34  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: NOVARTIS Corporation  
 STREET: 564 Morris Avenue  
 CITY: Summit  
 STATE: New Jersey  
 COUNTRY: USA  
 ZIP: 07901-6940  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/465,473B  
 FILING DATE: 5 June 1995  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/828,832  
 FILING DATE: 31-JAN-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: GB 91-810079.3  
 FILING DATE: 05-FEB-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Pfeiffer, Henna J.  
 REGISTRATION NUMBER: 22,640  
 REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (908)522 6940  
 TELEFAX: (908)522 6955  
 INFORMATION FOR SEQ ID NO: 14:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 637 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-465-473B-14

Query Match 53.5%; Score 1307; DB 2; Length 637;  
 Best Local Similarity 98.0%; Pred. No. 3, 1e-100;  
 Matches 243; Conservative 2; Mismatches 3; Indels 0; Gaps 0;

QY 15 KVPKQIQVQSGPPELKKGEYTKISCKASGYPTNYGNMWWYQAGGGLKMMGWINTSTG 74  
 DB 29 KLASVQVQSGPELKKGEYTKISCKASGYPTNYGNMWWYQAGGGLKMMGWINTSTG 88  
 QY 75 ESTPADDFKGRDFSLSTANTAYVLIQINNLSKSEDMATYFCARWEVYHGVYPMGGTTVT 134  
 DB 89 ESTPADDFKGRDFSLSTANTAYVLIQINNLSKSEDMATYFCARWEVYHGVYPMGGTTVT 148

QY 135 VSSGGGSGGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVNANVAMVQOKPQ 194  
 DB 149 VSSGGGSGGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVNANVAMVQOKPQ 208  
 QY 195 SPKLIYASASSRYTGVPSRFTGSGSPDFTTISVQAEADLAVVFCQOHFRTPTFGSGT 254  
 DB 209 SPKLIYASASSRYTGVPSRFTGSGSPDFTTISVQAEADLAVVFCQOHFRTPTFGSGT 268  
 QY 255 KLEIKALE 262  
 DB 269 KLEIKALE 276

RESULT 5  
 US-08-235-838-7  
 ; Sequence 7, Application US/08235838  
 ; Patent No. 5571894

GENERAL INFORMATION:  
 APPLICANT: Weis, Winfried S.  
 APPLICANT: Hynes, Nancy E.  
 APPLICANT: Harwerth, Ina-Maria  
 APPLICANT: Groner, Bernd  
 APPLICANT: Hardman, No. 5571894man  
 APPLICANT: Zwickl, Markus  
 TITLE OF INVENTION: Recombinant Antibodies Specific for a  
 TITLE OF INVENTION: Growth Factor Receptor  
 NUMBER OF SEQUENCES: 16  
 CORRESPONDENCE ADDRESSES:  
 ADDRESSEE: CIBA-GEIGY Corporation  
 STREET: 7 Skyline Drive  
 CITY: Hawthorne  
 STATE: New York  
 COUNTRY: USA  
 ZIP: 10532  
 COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent in Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/235,838  
 FILING DATE: TBA  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 07/828,832  
 FILING DATE: 31-JAN-1992  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: GB 91-810079.3  
 FILING DATE: 05-FEB-1991  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Elmer, James Scott  
 REGISTRATION NUMBER: 36,129  
 REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: (919)541-8614  
 TELEFAX: (919)541-8689  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 711 amino acids  
 TYPE: amino acid  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 US-08-235-838-7

Query Match 53.4%; Score 1303; DB 1; Length 711;  
 Best Local Similarity 94.2%; Pred. No. 7, 8e-100;  
 Matches 245; Conservative 2; Mismatches 1; Indels 12; Gaps 1;

QY 19 QIQVQSGPELKKGEYTKISCKASGYPTNYGNMWWYQAGGGLKMMGWINTSTGSESTF 78  
 DB 23 QIQVQSGPELKKGEYTKISCKASGYPTNYGNMWWYQAGGGLKMMGWINTSTGSESTF 82  
 QY 79 ADPDFKGRDFSLSTANTAYVLIQINNLSKSEDMATYFCARWEVYHGVYPMGGTTVTYSSG 138

Db	83	ADGFKRFPDPSLETSTANTAYLQINMLKSDMATYFCAMEYVHGVPWGGTITVYSSG	142
Qy	139	GGGGGGGGGGSSDIQLTQSHKFLSTSGDVSLTCAASODVYNAAVYQKFGSPKL	199
Db	143	GGGGGGGGGGSSDIQLTQSHKFLSTSGDVSLTCAASODVYNAAVYQKFGSPKL	202
Qy	199	LIYSASRRYTVGPSREFTSGSGPDFTFTSISSVQAEIDLAVYFCQOHFRTPFPFGSGTLEI	258
Db	203	LIYSASRRYTVGPSREFTSGSGPDFTFTSISSVQAEIDLAVYFCQOHFRTPFPFGSGTLEI	262
Qy	259	KALEISNSVWFSSVVPVLT	278
Db	263	KALE-----PVLE	270

```

RESULT 6
US-08-465-473B-7
: Sequence 7, Application US/08465473B
: Patent No. 5939531
:
: GENERAL INFORMATION:
: APPLICANT: Wels, Winfried S.
: APPLICANT: Hynes, Nancy E.
: APPLICANT: Harwerth, Ina-Maria
: APPLICANT: Groner, Bernd
: APPLICANT: Hardman, No. 5939531man
:
: TITLE OF INVENTION: Recombinant Antibodies Specific for a
: TITLE OF INVENTION: Growth Factor Receptor
: NUMBER OF SEQUENCES: 34
:
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: NOVARTIS Corporation
: STREET: 564 Morris Avenue
: CITY: Summit
: STATE: New Jersey
: COUNTRY: USA
: ZIP: 07901-6940
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patentin Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/465,473B
: FILING DATE: 5 June 1995
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: US 07/828,832
: FILING DATE: 31-JAN-1992
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: GB 91-810079.3
: FILING DATE: 05-FEB-1991
: ATTORNEY/AGENT INFORMATION:
: NAME: Pfeiffer, Henna J.
: REGISTRATION NUMBER: 22,640
: REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (908)522 6940
: TELEFAX: (908)522 6955
: INFORMATION FOR SEQ ID NO: 7:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 711 amino acids
: TYPE: amino acid
: TOPOLOGY: linear
: MOLECULE TYPE: protein
:
: US-08-465-473B-7

```

	Query Match	53 4*	Score 1303;	DB 2,	Length 711;
	Best Local Similarity	94.2%	Pred.	No. 7..8e-100;	
	Matches 245; Conservative	2;	Mismatches 1;	Indels 12;	Gaps
Oy	19 QIQIVDSGPELKKRGFTVTKISCKASAGPPFNNGMMNVKAPQGLKRMGININTSGESLF	78	:	:	:: ::

Db	23	QVQLQQSGFELKPKRGFTATISCKASQYPTNYGMNWVQAQPGGLKMMGWINTSGSTF	82
QY	79	ADDPKRPDPSLETANTAYLIQINLKSDMATYTCAMETVHGVPWGGGTTVWSSG	138
Db	83	ADDPKRPDPSLETISANTAYLIQINLKSDMATYTCAMEVYHGVPWGGGTTVWSSG	142
QY	139	GGSGGGSGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNAVMVYQKRGOSP	198
Db	143	GGSGGGSGSGGGSDIQLTQSHKFLSTSGDRVSIITCKASODVYNAVMVYQKRGOSP	202
QY	199	LIYSASSRTTGVPSRRTSGSGGPDFTTISVQAEADLAVYFCQHFRTPTFGSGTLEI	258
Db	203	LIYSASSRRTGVPSRRTSGSGGPDFTTISVQAEADLAVYFCQHFRTPTFGSGTLEI	262
QY	259	KALEISNSVMYFSSVYVPLQ	278
Db	263	KALE-----VPLQ	270

```

      RESULT 7
US-08-840-713-2
; Sequence 2, Application US/08840713
; Patent No. 6498233
; GENERAL INFORMATION:
; APPLICANT: WELLS, Winfried, Dr.
; APPLICANT: FOYMINAVA, Jesus
; TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP
; STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
; City: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-5701
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,713
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kiltz, Monica Chin
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: 1614-7014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638 - 5000
; TELEFAX: (202) 638 - 4810
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 530 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
US-08-840-713-2

```

Query Match	53.3%	Score 1302	DB 4	Length 530
Best Local Similarity	99.2%	Pred. No. 6	3e-100	
Matches	242	Conservative	1	Mismatches 1; Indels 0; Gaps 0
QY	19	QIQLVQSSPELKKGEIVTKISCRASGYPTFNYYGMNWKAPGGGLKMGWINTSGESTF	78	
	:	:	:	
Db	18	QVQLQDSPELKKGEIVTKISCRASGYPFNYYGMNWKAPGGGLKMGWINTSGESTF	77	
QY	79	ADDEKGEPDFLEFSANTAYLIQINNLKSEDMATYFCARBEVYHGYPYVGQGTYYVSSG	138	
Db	78	ADPEKGRDPLEFSANTAYLIQINNLKSEDMATYFCARBEVYHGYPYVGQGTYYVSSG	137	
QY	139	GGSGGGGGGGGGSDIQLTQSHKPLTSTVCDRVSTTCASQDVVNAVAVYQQKPGQSPKL	198	

Db 138 GGGSGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNNAVWQOKPGQSPKL 197  
Qy 199 LIYSASSRYTGVPSRFITGSGGSDPFTTISVQAEADLAVYFCQGHFRTPTFGSGTKLEI 258  
Db 198 LIYSASSRYTGVPSRFITGSGGSDPFTTISVQAEADLAVYFCQGHFRTPTFGSGTKLEI 257  
Qy 259 KALE 262  
Db 258 KALE 261

## RESULT 8

US-08-235-838-5  
; Sequence 5, Application US/08235838  
; Patent No. 5571894  
; GENERAL INFORMATION:  
; APPLICANT: Wels, Winfried S.  
; APPLICANT: Hynes, Nancy E.  
; APPLICANT: Harwerth, Ina-Maria  
; APPLICANT: Groner, Bernd  
; APPLICANT: Hardman, No. 5571894man  
; APPLICANT: Zwickl, Markus  
; TITLE OF INVENTION: Recombinant Antibodies Specific for a  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10512  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/235.838  
; FILING DATE: TBA  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/828,832  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 91-810079.3  
; FILING DATE: 05-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elmer, James Scott  
; REGISTRATION NUMBER: 36,129  
; REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919)541-8614  
; TELEFAX: (919)541-8689  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 241 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-235-838-5

Query Match 52.6%; Score 1284; DB 1; Length 241;  
Best Local Similarity 99.2%; Pred. No. 6.9e-99;  
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 19 QIOLVOSGPELKKRPETVKISCKASGYPTNYGNMWWQAPOGGLKMMGWINTSGESTF 78  
Db 2 QVQLQQSGPELKKRPETVKISCKASGYPTNYGNMWWQAPOGGLKMMGWINTSGESTF 61  
Qy 79 AADFKGRDFSLFETISANTAYLQINNLKSEDMATYFCARMEVYHGVVPYWGQTTVTYSSG 138  
Db 62 AADFKGRDFSLFETISANTAYLQINNLKSEDMATYFCARMEVYHGVVPYWGQTTVTYSSG 121

Qy 139 GGGSGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNNAVWQOKPGQSPKL 198  
Db 122 GGGSGGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASODVYNNAVWQOKPGQSPKL 181  
Qy 199 LIYSASSRYTGVPSRFITGSGGSDPFTTISVQAEADLAVYFCQGHFRTPTFGSGTKLEI 258  
Db 182 LIYSASSRYTGVPSRFITGSGGSDPFTTISVQAEADLAVYFCQGHFRTPTFGSGTKLEI 241

## RESULT 9

US-08-465-473B-5  
; Sequence 5, Application US/08465473B  
; Patent No. 5939531  
; GENERAL INFORMATION:  
; APPLICANT: Wels, Winfried S.  
; APPLICANT: Hynes, Nancy E.  
; APPLICANT: Harwerth, Ina-Maria  
; APPLICANT: Groner, Bernd  
; APPLICANT: Hardman, No. 5939531man  
; APPLICANT: Zwickl, Markus  
; TITLE OF INVENTION: Recombinant Antibodies Specific for a  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NOVARTIS Corporation  
; STREET: 564 Morris Avenue  
; CITY: Summit  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07901-6940  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465.473B  
; FILING DATE: 5 June 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/828,832  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 91-810079.3  
; FILING DATE: 05-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pfeiffer, Henna J.  
; REGISTRATION NUMBER: 22,640  
; REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)522 6940  
; TELEFAX: (908)522 6955  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 241 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-465-473B-5

Query Match 52.6%; Score 1284; DB 2; Length 241;  
Best Local Similarity 99.2%; Pred. No. 6.9e-99;  
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 19 QIOLVOSGPELKKRPETVKISCKASGYPTNYGNMWWQAPOGGLKMMGWINTSGESTF 78  
Db 2 QVQLQQSGPELKKRPETVKISCKASGYPTNYGNMWWQAPOGGLKMMGWINTSGESTF 61  
Qy 79 AADFKGRDFSLFETISANTAYLQINNLKSEDMATYFCARMEVYHGVVPYWGQTTVTYSSG 138  
Db 62 AADFKGRDFSLFETISANTAYLQINNLKSEDMATYFCARMEVYHGVVPYWGQTTVTYSSG 121

QY 139 GGGGGGGGGGGGGSDIQLTQSHKFLSTSVGDVSTCTCASGDVYNAVMYQKRGQSFKL 198

Db 122 GGGGGGGGGGGSDIQLTQSHKFLSTSVGDVSTCTCASGDVYNAVMYQKRGQSFKL 181

QY 199 LITASASRTGVSRRTGSGSGSPDPFTTSSVQADLAVYCCOHPRPPTFGSGTKLET 258

Db 182 LITASASRTGVSRRTGSGSGDPDFTTSSVQADLAVYCCOHPRPPTFGSGTKLET 241

RESULT 10

Sequence 2, Application US/08133804  
Patent No. 5534254  
GENERAL INFORMATION:  
APPLICANT: Huston, James S.  
APPLICANT: Oppermann, Hermann  
APPLICANT: Houston, L. L.  
APPLICANT: Ring, David B.  
TITLE OF INVENTION: Biosynthetic Binding Proteins For  
NUMBER OF SEQUENCES: 11  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Testa, Horwitz & Thibault/Patent Department  
STREET: Exchange Place, 53 State Street  
CITY: Boston  
STATE: Massachusetts  
COUNTRY: USA  
ZIP: 02109  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/133,804  
FILING DATE:  
CLASSIFICATION: 424  
ATTORNEY/AGENT INFORMATION:  
NAME: Kelley, Robin D.  
REGISTRATION NUMBER: 34,637  
REFERENCE/DOCKET NUMBER: 2054/22  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617-248-7477  
TELEFAX: 617-248-7100  
INFORMATION FOR SEQ ID NO: 2:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 250 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-133-804-2

Query Match	40.2%;	Score 981.5;	DB 1;	Length 250;
Best Local Similarity	76.8%;	Pred. No. 1e-73;		
Matches 185; Conservative	22;	Mismatches 33;	Indels 1;	Gaps 1;

QY	19	QIQLVOSGPELKKPGEYTKISCSKAGVPEFTNYGNNWYKQAPGGLKMMGIMTNSGESEF	78
QY	3	EIQILOSGBELKKPGEYTKISCSKSGTFTNYGNNWYKQAPGGLKMMGIMTNSGESEY	62
QY	79	ADDFKGRDPFSLSTSSANTAYLQIINNLKSEDMATFCARMEVYHGYPVYGGGTYYVS	138
QY	63	AEEFGKRAFSLETSSASTAYLQIINNLKNEPATYFQCRQFTITYGFRANNGQGLIYVSA-	121
QY	139	GGSGSGGGSGGGGSGIQLTOSHKFLSTSVGDRVSYITKASQDYVNAVANYOQFPGSPKL	198
QY	122	SSSSSSSSSSSSSIIVMTOSPKMSTSVSGRVAISCSKASQDVSTVANYOQFPGSPKL	181
QY	199	LIYSASRYTGVPSRPFSTSGSGDPFTFISVSOEADLAVYFCQOHRPFTFSGSGKLEI	258
QY	162	LIYMTSTHTGVPRDFRSTSGSGTDTYTLTISVQAEADLALHYCQOHRKVPFTFSGSGKLEI	241
QY	259	K 259	

Db 242 K 242

## RESULT 11

US-08-461-184-8  
; Sequence 8, Application US/08461184  
Patent No. 7,533,170

GENERAL INFORMATION:

APPLICANT: DORAI, HAIMANAI  
APPLICANT: OPPERMAN, HERRANN  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN  
TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA  
NUMBER OF SEQUENCES: 10

ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 01948

```

? COMPUTER READABLE FORM:
? MEDIUM TYPE: Floppy disk
? COMPUTER: IBM PC compatible
? OPERATING SYSTEM: PC-DOS/MS-DOS
? SOFTWARE: Patent in Release #1.0, Version #1.25
? CURRENT APPLICATION DATA:

```

FILING DATE:

; PRIOR APPLICATION DATA:

FILING DATE: \_\_\_\_\_

NAME: KELLEY, ROBIN D

REFERENCE/DOCKET NUMBER: CRP0933

TELEPHONE: 617/248-7000

; INFORMATION FOR SEQ ID NO: 8  
SEQUENCE AND ORIGIN  
ORIGIN

LENGTH: 250 amino acids

TOPOLOGY: linear

US-08-461-184-8

Query Match	40
Post-Topic Classification	76

Malchus 185; Conservativ

09 19 0106VQSGPELKKF

DD 3 E19LVQS6PELWNE

19 ADDRESSES LISTED

DD 63 AEEFGRAFSLEI

[illegible][illegible]

133 L110H00K11GVFS  
27

202 D11M101110VED

CCZ Y CCZ

23 717 N 747



```

; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESS: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,838
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-838-2

Query Match      40.2%; Score 981.5; DB 1; Length 250;
Best Local Similarity 76.8%; Pred. No. 1e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 19 QIOLVSGPELKKRGETVKISKASGYPTNYGMNVKAPGQGLKMMGWINTSGESTF 78
Db 3 EIQLVSGPELKKRGETVKISKASGYPTNYGMNVKAPGQGLKMMGWINTSGESTF 62
QY 79 ADFKGRFDSLETSANTAYLIQINNLSKSDMATYFCARMEVYHGVYPYWGQGTIVVSSG 138
Db 63 AEEFGRAFSLFETASATAYLIQINNLSKSDMATYFCARMEVYHGVYPYWGQGTIVVSSG 121
QY 139 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSTICKASQDVYNAVAMVYQKPGQSPKL 198
Db 122 SSSSGSSSSSSSDIQLTQSHKFLSTSVGDRVSTICKASQDVYNAVAMVYQKPGQSPKL 181
QY 199 LIYSASRYTGVPDRFTGSGGPDFTFTISSVOAEDLAVYFCQGHFRPTFGSGTKLEI 258
Db 182 LIYMTSRHTGVDRFTGSGGPDFTFTISSVOAEDLAVYFCQGHFRPTFGSGTKLEI 241
QY 259 K 259
Db 242 K 242

RESULT 15
; US-08-461-386-2
; Sequence 2, Application US/08461386
; Patent No. 5837846
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
```

```

; ADDRESS: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,386
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-386-2

Query Match      40.2%; Score 981.5; DB 2; Length 250;
Best Local Similarity 76.8%; Pred. No. 1e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 19 QIOLVSGPELKKRGETVKISKASGYPTNYGMNVKAPGQGLKMMGWINTSGESTF 78
Db 3 EIQLVSGPELKKRGETVKISKASGYPTNYGMNVKAPGQGLKMMGWINTSGESTF 62
QY 79 ADFKGRFDSLETSANTAYLIQINNLSKSDMATYFCARMEVYHGVYPYWGQGTIVVSSG 138
Db 63 AEEFGRAFSLFETASATAYLIQINNLSKSDMATYFCARMEVYHGVYPYWGQGTIVVSSG 121
QY 139 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSTICKASQDVYNAVAMVYQKPGQSPKL 198
Db 122 SSSSGSSSSSSSDIQLTQSHKFLSTSVGDRVSTICKASQDVYNAVAMVYQKPGQSPKL 181
QY 199 LIYSASRYTGVPDRFTGSGGPDFTFTISSVOAEDLAVYFCQGHFRPTFGSGTKLEI 258
Db 182 LIYMTSRHTGVDRFTGSGGPDFTFTISSVOAEDLAVYFCQGHFRPTFGSGTKLEI 241
QY 259 K 259
Db 242 K 242
```

Search completed: April 20, 2005, 07:08:35  
Job time : 21.9082 secs





Db 141 SSGGGGGGGGGGGSDIOMTQSSSSFSVSLGDRVITTCANEDINNRLAMVQQTGNS 200  
QY 136 PTLIIYASASRTYGVPSRTGSGSGDPFTTISVQAEPLANYFCQGHRTPTTGSSTK 255  
Db 201 PRLISGATNLVTVPSRFSGSGSGKDYTLITTSLOAEDFATYCCQWSTFTTGSSTGTE 260  
QY 256 LEIK-----ALEISNVWFSSVVLQ-- 278  
Db 261 LEIKVPEKSSDKTHCPCEPAPELLGGSVFLPPKPKDMLMISRTPEVTCVVDVSHED 320  
QY 279 -----KVNSTTKP-----VLR-----TPS 293  
Db 321 PEKFNWYVDVGVENAKTKPREEQYNSTYRVSVLTVLHQDMLNGEKYCKVSKALPA 380  
QY 294 PVHPT---GTSQPPPE-----DCEPPGS 314  
Db 361 PLEKTSIAKQGPPEPOVYTLPPSRDELTKQVSLTCLVKGFPSPDAVWESNQPPNN 440  
QY 315 VKGTG--LDPLEDPKLYLDGILFLY----- 339  
Db 441 YKTPPVLD-----SDGSFLYKSLTVDKSRMOQGVNFSGVNHEALHNYTOKS 490  
QY 340 -----GVII---TALYLRAKFSRSAAETANLQDPNOLYNELNGRREE 379  
Db 491 LSLSPKMALIYLGAVAGLLFLIGIFFRVYKFSRSDADAPAYQOGQNLVNEMLNGRREE 550  
QY 360 YDVELEKKARPDEMGKQOORRNPOEGVYNALQDKMAEAYSEITGKERRRGKHGDLX 439  
Db 551 YDVLDRKRGPRDEMGK--PRRKNPQEGLYNELQDKMAEAYSEIGMKERRRGKHGDLX 609  
QY 440 QGLSTATDYDALHMOTLAPR 461  
Db 610 QGLSTATDYDALHMOTLAPR 631

## RESULT 2

US-10-006-773-2  
; Sequence 2, Application US/10006773  
; Publication No. US20020132983A1  
; GENERAL INFORMATION:  
; APPLICANT: Jungmans, Richard P.  
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti-  
; FILE REFERENCE: 003  
; CURRENT APPLICATION NUMBER: US/10/006,773  
; CURRENT FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: 60/250,089  
; PRIOR FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 443  
; TYPE: PRT  
; ORGANISM: Homo sapiens and Mus sp.  
US-10-006-773-2

Query Match. 41.6%; Score 1015.5; DB 13; Length 443;  
Best Local Similarity 52.0%; Pred. No. 1e-62;  
Matches 225; Conservative 25; Mismatches 58; Indels 125; Gaps 8;

QY 149 GGGSDIQLTQSHKFLTSVGDVRSITCKASODVYNNAVAMYOQKPGSPKLIYSASRYT 208  
Db 16 GVHSDIQLTQSPSSISASVGDVITITCKASODVGTSVAMYOQKPKAKLIIWTSTRHT 75  
QY 209 GVPFRFTGSGGPDFTTISVQAEPLAVYFCQGH--FRTPFTGSGTKLEIK----- 259  
Db 76 GVPFRFTGSGGPDFTTISVQAEPLAVYFCQGH--FRTPFTGSGTKLEIK----- 259  
QY 260 -----ALEISNV----- 267  
Db 133 GSGSGSGSEVQLVESGGGVVQPGRLRLSCSASGPDFTTYMMSWVQAQPKGLEWIGEI 192  
QY 268 -----MYFSS 272

Db 193 HPDSSTINVASLDRFTISRDNAKNTLFLQMDSLRPEDTCVYFCASLYGFPWFAYWQ 252  
QY 273 VVPLQKVNSTTKKPVLRTPSPVHPGTGSP--QRPDCRPR--GSYKGTGLDFLEDPKL 328  
Db 253 GTPVTVSAAKPTTTPADRPPTPA--PTIASQPLSRPAAAPAGAVHTRGDLPALDKL 311  
QY 329 CYLLDGLIFLYGVTITLALYLRKFSRSAAETANLQDPNOLYNELNGRREEVDYLEKKA 388  
Db 312 CYLLDGLIFLYGVTITLALYLRKFSRSAAETANLQDPNOLYNELNGRREEVDYLEKKA 388  
QY 389 RDPENMGKQOORRNPOEGVYNALQDKMAEAYSEITGKERRRGKHGDLXQGLSTATKD 448  
Db 372 RDPENMGK--PRRKNPQEGLYNELQDKMAEAYSEIGMKERRRGKHGDLXQGLSTATKD 430  
QY 449 TYDALHMOTLAPR 461  
Db 431 TYDALHMOTLAPR 443

## RESULT 3

US-10-006-771A-2  
; Sequence 2, Application US/10006771A  
; Publication No. US20020165360A1  
; GENERAL INFORMATION:  
; APPLICANT: Jungmans, Richard P.  
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen  
; FILE REFERENCE: 002  
; CURRENT APPLICATION NUMBER: US/10/006,771A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: 60/250,090  
; PRIOR FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 2  
; LENGTH: 443  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
US-10-006-771A-2

Query Match 41.6%; Score 1015.5; DB 13; Length 443;  
Best Local Similarity 52.0%; Pred. No. 1e-62;  
Matches 225; Conservative 25; Mismatches 58; Indels 125; Gaps 8;

QY 149 GGGSDIQLTQSHKFLTSVGDVRSITCKASODVYNNAVAMYOQKPGSPKLIYSASRYT 208  
Db 16 GVHSDIQLTQSPSSISASVGDVITITCKASODVGTSVAMYOQKPKAKLIIWTSTRHT 75  
QY 209 GVPFRFTGSGGPDFTTISVQAEPLAVYFCQGH--FRTPFTGSGTKLEIK----- 259  
Db 76 GVPFRFTGSGGPDFTTISVQAEPLAVYFCQGH--FRTPFTGSGTKLEIK----- 259  
QY 260 -----ALEISNV----- 267  
Db 133 GSGSGSGSEVQLVESGGGVVQPGRLRLSCSASGPDFTTYMMSWVQAQPKGLEWIGEI 192  
QY 268 -----MYFSS 272  
Db 193 HPDSSTINVASLDRFTISRDNAKNTLFLQMDSLRPEDTCVYFCASLYGFPWFAYWQ 252  
QY 273 VVPLQKVNSTTKKPVLRTPSPVHPGTGSP--QRPDCRPR--GSYKGTGLDFLEDPKL 328  
Db 253 GTPVTVSAAKPTTTPADRPPTPA--PTIASQPLSRPAAAPAGAVHTRGDLPALDKL 311  
QY 329 CYLLDGLIFLYGVTITLALYLRKFSRSAAETANLQDPNOLYNELNGRREEVDYLEKKA 388  
Db 312 CYLLDGLIFLYGVTITLALYLRKFSRSAAETANLQDPNOLYNELNGRREEVDYLEKKA 388  
QY 389 RDPENMGKQOORRNPOEGVYNALQDKMAEAYSEITGKERRRGKHGDLXQGLSTATKD 448  
Db 372 RDPENMGK--PRRKNPQEGLYNELQDKMAEAYSEIGMKERRRGKHGDLXQGLSTATKD 430  
QY 449 TYDALHMOTLAPR 461



```

/ MEDIUM TYPE: Diskette
/ COMPUTER: IBM Compatible
/ OPERATING SYSTEM: DOS
/ SOFTWARE: FASTSEQ for Windows Version 2.0
/ CURRENT APPLICATION DATA:
/   APPLICATION NUMBER: US/08/812.393A
/   FILING DATE: 05-MAR-1997
/   CLASSIFICATION: 435
/   PRIOR APPLICATION DATA:
/     APPLICATION NUMBER:
/     FILING DATE:
/   ATTORNEY/AGENT INFORMATION:
/     NAME: Murashige, Kate H
/     REGISTRATION NUMBER: 29,959
/     REFERENCE/DOCKET NUMBER: 31333-20001.00
/   TELECOMMUNICATION INFORMATION:
/     TELEPHONE: 202-887-1500
/     TELEFAX: 202-882-0168
/   TELEX:
/   INFORMATION FOR SEQ ID NO: 2:
/     SEQUENCE CHARACTERISTICS:
/       LENGTH: 444 amino acids
/       TYPE: amino acid
/       STRANDEDNESS: single
/       TOPOLOGY: linear
/     MOLECULE TYPE: protein
/     FRAGMENT TYPE: internal
/   US-08-812-393A-2

```

```

Query Match      38.1%; Score 930; DB 8; Length 444;
Best Local Similarity 48.2%; Pred. No. 9.4e-57;
Matches 217; Conservative 50; Mismatches 145; Indels 38; Gaps 14;

```

```

QY 21 QLVSGPE--LKKPGETVTKISCKASGYPTNYGNMNVKQAPGGGLKMKMGINTSTGESETF 78
DB 24 QVVOQSPASLVLDGEENAELOCSFS--IFTNQ-VQMFYORGGRLVSLVNPSTGKOS-- 78
QY 79 ADDEKGRFDFSLSEANTAYLIQINLKSMDMATYFCARMEVYHGVPYWGQGTVTYVSSG 138
DB 79 -----GRLTSTTVYIKERRSLHSSQITDSGTIYCANSSGSAKLTFKGTGLSVKSG 133
QY 139 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNAVAMYQKPGQSPKL 198
DB 134 GGGSGGGSGGGGSEAAVTQSPRNKVAVTGKVTLSQNTNN--HNNMYWYRQDTGHLRL 192
QY 199 LIYS---ASSRYTGVPSPRFTGS--GSGPDTFTISSVQAEIAYVFC---QQHRTPTFTG 251
DB 193 IHTSYGAGSTKEDIPDGYKASRPSQENFSLVELATPSTQTSVYFCASGETGTNERLFG 252
QY 252 SGTLEIKALEISNSVMYFSSVVPVLOKVNSTTKKVLRTSPVPVHPTGTSQORPEDCRP 311
DB 253 HGTKLSTVLT---SNSIMYFHFVFLPAKPTTT--PAPRPPTPA--PTIASQPL---SLRP 304
QY 312 RGSVKGITGLDLEDPKLCYLLDGLIFTYGVITLALYLAKFSASAEFTANLQDPNOLYNE 371
DB 305 SSS-----RDPKLCYLLDGLIFTYGVITLALFLRVKFSRSADAPAYQCGQONOLYNE 355
QY 372 LNLGRREYDVLEKKRAPDEMGKQOORRNPOEGVYNALQKDMAEYSEIGTKGERRR 431
DB 356 LNLGRREYDVLDKRGKRDPEMGK--PRKNPQEGVLNELQKDMAEYSEIGTKGERRR 414
QY 432 GKGHGGLYQGLSTATKDTYDALHMQTLAPR 461
DB 415 GKGHGGLYQGLSTATKDTYDALHMQALPPR 444

```

```

RESULT 7
US-09-774-681-2
/ Sequence 2, Application US/09774681
/ Publication No. US20030208780A1
/ GENERAL INFORMATION:
/   APPLICANT: Sunol Molecular Corporation
/   APPLICANT: Sherman, Linda

```

```

/ APPLICANT: Luetgarten, Joseph
/ TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL
/ TITLE OF INVENTION: RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR ANTIGENS
/ FILE REFERENCE: 31333-20001.01
/ CURRENT APPLICATION NUMBER: US/09/774.681
/ PRIOR FILING DATE: 2001-02-01
/ PRIOR APPLICATION NUMBER: US 08/812.393
/ PRIOR FILING DATE: 1997-03-05
/ PRIOR APPLICATION NUMBER: US 60/012,845
/ PRIOR FILING DATE: 1996-03-05
/ NUMBER OF SEQ ID NOS: 65
/ SOFTWARE: FASTSEQ for Windows Version 4.0
/ SEQ ID NO: 2
/ LENGTH: 449
/ TYPE: PRT
/ ORGANISM: Artificial Sequence
/ FEATURE:
/ OTHER INFORMATION: Deduced amino acid derivative of effective T cell
/ OTHER INFORMATION: receptor
/ US-09-774-681-2

```

```

Query Match      38.0%; Score 923; DB 10; Length 449;
Best Local Similarity 48.2%; Pred. No. 1.3e-56;
Matches 217; Conservative 50; Mismatches 145; Indels 38; Gaps 14;

```

```

QY 21 QLVSGPE--LKKPGETVTKISCKASGYPTNYGNMNVKQAPGGGLKMKMGINTSTGESETF 78
DB 24 QVVOQSPASLVLDGEENAELOCSFS--IFTNQ-VQMFYORGGRLVSLVNPSTGKOS-- 78
QY 79 ADDEKGRFDFSLSEANTAYLIQINLKSMDMATYFCARMEVYHGVPYWGQGTVTYVSSG 138
DB 79 -----GRLTSTTVYIKERRSLHSSQITDSGTIYCANSSGSAKLTFKGTGLSVKSG 133
QY 139 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSTTCASQDVYNAVAMYQKPGQSPKL 198
DB 134 GGGSGGGSGGGGSEAAVTQSPRNKVAVTGKVTLSQNTNN--HNNMYWYRQDTGHLRL 192
QY 199 LIYS---ASSRYTGVPSPRFTGS--GSGPDTFTISSVQAEIAYVFC---QQHRTPTFTG 251
DB 193 IHTSYGAGSTKEDIPDGYKASRPSQENFSLVELATPSTQTSVYFCASGETGTNERLFG 252
QY 252 SGTLEIKALEISNSVMYFSSVVPVLOKVNSTTKKVLRTSPVPVHPTGTSQORPEDCRP 311
DB 253 HGTKLSTVLT---SNSIMYFHFVFLPAKPTTT--PAPRPPTPA--PTIASQPL---SLRP 304
QY 312 RGSVKGITGLDLEDPKLCYLLDGLIFTYGVITLALYLAKFSASAEFTANLQDPNOLYNE 371
DB 305 SSS-----RDPKLCYLLDGLIFTYGVITLALFLRVKFSRSADAPAYQCGQONOLYNE 355
QY 372 LNLGRREYDVLEKKRAPDEMGKQOORRNPOEGVYNALQKDMAEYSEIGTKGERRR 431
DB 356 LNLGRREYDVLDKRGKRDPEMGK--PRKNPQEGVLNELQKDMAEYSEIGTKGERRR 414
QY 432 GKGHGGLYQGLSTATKDTYDALHMQTLAPR 461
DB 415 GKGHGGLYQGLSTATKDTYDALHMQALPPR 444

```

```

RESULT 8
US-10-239-656-79
/ Sequence 79, Application US/10239656
/ Publication No. US2004003839A1
/ GENERAL INFORMATION:
/   APPLICANT: KUPER, PETER
/   APPLICANT: RIETHMULLER, GERT
/   APPLICANT: LUTTERBUSE, RALE
/   APPLICANT: BORSCHERT, KATRIN
/   APPLICANT: KISCHEL, ROMAN
/   APPLICANT: MAYER, MONIKA
/   APPLICANT: HOFMEISTER, ROBERT
/ TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE
/ TITLE OF INVENTION: TO AN EPIPOPE OF THE NKGD2 RECEPTOR COMPLEX
/ FILE REFERENCE: 029976/0106

```

```

; CURRENT APPLICATION NUMBER: US/10/239,656
; CURRENT FILING DATE: 2003-03-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03414
; PRIOR FILING DATE: 2001-03-26
; PRIOR APPLICATION NUMBER: EP 00106467.4
; PRIOR FILING DATE: 2000-03-24
; NUMBER OF SEQ ID NOS: 92
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 79
; LENGTH: 505
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic 3810xP5-
; OTHER INFORMATION: 23 bispecific single chain Fv
US-10-239-656-79

```

```

Query Match      37.6%; Score 917.5; DB 15; Length 505;
Best Local Similarity 50.5%; Pred. No. 8,1e-56;
Matches 185; Conservative 25; Mismatches 31; Indels 125; Gaps 4;

```

```

QY 19 QIOLVSGPELKKPEETVYKISCKASGYPTNYGMNVAQAPOGGLKMMGMINTSGESTF 78
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 128 QVQLQDSGPPELKKPEETVYKISCKASGYFTNYGMNVAQAPEGKRMGMINTYGEPTY 187
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 79 ADFKGRFDFSLSTANTAYVLIQINNLKSEDMATYFCAR----- 116
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 188 GDFKGRFAFSLSTASTAYVLIQINNLKEDTATYFCARTSPDYGGCTTVTVSSGGGS 247
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 117 -----W----- 117
Db 248 EVQLLESGGGLVQPGSGSLKLSCAASGDFPSRYMWSMVRQAPGKLEIMGEINPDSSTINX 307
QY 118 -----EYH-----GYTPYGGCTTVTVSSGG 139
Db 308 TPSLKRPPIISRDNAKNTLYIQMSKVRSEDTALYYCARLGQWGYDYGQGTTVTVSSGG 367
QY 140 GSGGGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNA-----VAWTQOKPG 193
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 368 GSGGGGGSGGGSELVMTQSPSSLTVTAGERVMTSCKSQSLNSGNQKNTITWYQOKPG 427
QY 194 QSPKLLIYSASSRYTVGPSRFTSGSGSDPTFTTISVQAEDLAVYFCQGHFTPTFTSGS 253
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 428 QPKLLIYMASTRREGVDPDRFTGSSGTDFTLTISVQAEDLAVYVCQNDYSYPLTFGAG 487
QY 254 TKLEIK 259
   |||:|||||
Db 488 TKLEIK 493

```

```

RESULT 9
US-09-766-543-10
; Sequence 10, Application US/09766543
; Patent No. US20020041865A1
; GENERAL INFORMATION:
; APPLICANT: Auslin, Richard
; APPLICANT: Kwok, Cheuk S.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: METHODS FOR TREATING TUMORS
; FILE REFERENCE: PP01679.002
; CURRENT APPLICATION NUMBER: US/09/766,543
; CURRENT FILING DATE: 2000-01-19
; PRIOR APPLICATION NUMBER: 60/177,258
; PRIOR FILING DATE: 2000-01-20
; NUMBER OF SEQ ID NOS: 14
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: 520C9
; OTHER INFORMATION: humanized single-chain antibody used in the

```

```

; OTHER INFORMATION: IL-2-antibody fusions
US-09-766-543-10

```

```

Query Match      37.4%; Score 913; DB 9; Length 267;
Best Local Similarity 68.4%; Pred. No. 8,1e-56;
Matches 167; Conservative 37; Mismatches 36; Indels 4; Gaps 2;

```

```

QY 19 QIOLVSGPELKKPEETVYKISCKASGYPTNYGMNVAQAPOGGLKMMGMINTSGESTF 78
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 23 EMQLVESPEYKRPASVYKVSCKASGYFTNYGMNVAQAPOGGLKMMGMINTYGGSTY 82
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
QY 79 ADFKGRFDFSLSTANTAYVLIQINNLKSEDMATYFCARMEYHGYVPYGGCTTVTVSSG 138
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 83 ADFKERVLTMTDTSTAYMDLRSLRSDPTAVVYCAR---RFGPA-YMGCGTLTVTVSSG 138
QY 139 GSGGGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVMAVQOKPGQPKL 198
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 139 GSGGGGGSGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVMAVQOKPGQPKL 198
QY 199 LIYSASSRYTVGPSRFTSGSGSDPTFTTISVQAEDLAVYFCQGHFTPTFTSGSTKLEI 258
   |||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
Db 199 LIYATSLDSGVPSRFTSGSGSDPTFTTISVQAEDLAVYFCQGHFTPTFTSGSTKLEI 258
QY 259 KALE 262
Db 259 KGSE 262

```

```

RESULT 10
US-10-127-890-148
; Sequence 148, Application US/10127890
; Publication No. US20030166196A1
; GENERAL INFORMATION:
; APPLICANT: Better, Marc D.
; Caroll, Stephen F.
; Studnicka, Gary M.
; TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating
; Proteins
; NUMBER OF SEQUENCES: 173
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Mcandrews, Held & Malloy, Ltd.
; STREET: 500 West Madison Street, 34th floor
; CITY: Chicago
; STATE: Illinois
; COUNTRY: USA
; ZIP: 60661
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/10/127,890
; FILING DATE: 23-Apr-2002
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/646,360
; FILING DATE: 13-MAY-1996
; APPLICATION NUMBER: PCT/US94/05348
; FILING DATE: 12-MAY-1994
; APPLICATION NUMBER: US 08/064,691
; FILING DATE: 12-MAY-1993
; APPLICATION NUMBER: US 07/998,430
; FILING DATE: 09-DEC-1992
; APPLICATION NUMBER: US 07/901,707
; FILING DATE: 19-JUN-1992
; APPLICATION NUMBER: US 07/787,567
; FILING DATE: 04-NOV-1991
; ATTORNEY/AGENT INFORMATION:
; NAME: McNicholas, Janet M.
; REGISTRATION NUMBER: 32,918
; REFERENCE/DOCKET NUMBER: 200-70.P4
; TELECOMMUNICATION INFORMATION:

```

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 148:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 148:  
US-10-127-890-148

Query Match 37.0%; Score 903.5; DB 14; Length 240;  
Best Local Similarity 69.3%; Pred. No. 3.3e-55;  
Matches 167; Conservative 32; Mismatches 41; Indels 1; Gaps 1;

QY 19 QIQVQSGPELKKPGEYTKISCKASGYPTNYGNMVAQAPGQGLKMMGMINSTGSESTF 78  
DB 1 EIQLVQSGGGLVKGPGSGVRSICASAGYPTNYGNMVAQAPGQGLKMMGMINHTGEPY 60  
QY 79 ADDEKGRFDESLFETSAVTAYLQINNLSKSDMATYFCARMEVYHGVYVWGQGTITVYSSG 138  
DB 61 ADSFKGRFTSLDSDKNTAYLQINSLRADTAVYFCR-RGYDMYFPVWGQGTITVYSSG 119  
QY 139 GGGSGGGSGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVNAVAVYQKPGQSPKL 198  
DB 120 GGGSGGGSGGGGSDIQLTQSPSSLASVGDVITTCRASQDINSYLSWFOQKRGKAPKT 179  
QY 199 LIYASASRYTVPSRFTTSSGSGPDFTTISVQAEDLAVYFCQHFRTPTFGSGTKLEI 258  
DB 180 LIYANRLSGVPSRFTTSSGSGPDFTTISVQAEDLAVYFCQHFRTPTFGSGTKLEI 239  
QY 259 K 259  
DB 240 K 240

RESULT 11  
US-10-717-243-148  
Sequence 148, Application US/10717243  
Publication No. US20050054835A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/717,243  
FILING DATE: 18-NOV-2003  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,765  
FILING DATE: 15-APR-1997  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 148:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 148:  
US-10-717-243-148

Query Match 37.0%; Score 903.5; DB 17; Length 240;  
Best Local Similarity 69.3%; Pred. No. 3.3e-55;  
Matches 167; Conservative 32; Mismatches 41; Indels 1; Gaps 1;

QY 19 QIQVQSGPELKKPGEYTKISCKASGYPTNYGNMVAQAPGQGLKMMGMINSTGSESTF 78  
DB 1 EIQLVQSGGGLVKGPGSGVRSICASAGYPTNYGNMVAQAPGQGLKMMGMINHTGEPY 60  
QY 79 ADDEKGRFDESLFETSAVTAYLQINNLSKSDMATYFCARMEVYHGVYVWGQGTITVYSSG 138  
DB 61 ADSFKGRFTSLDSDKNTAYLQINSLRADTAVYFCR-RGYDMYFPVWGQGTITVYSSG 119  
QY 139 GGGSGGGSGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVNAVAVYQKPGQSPKL 198  
DB 120 GGGSGGGSGGGGSDIQLTQSPSSLASVGDVITTCRASQDINSYLSWFOQKRGKAPKT 179  
QY 199 LIYASASRYTVPSRFTTSSGSGPDFTTISVQAEDLAVYFCQHFRTPTFGSGTKLEI 258  
DB 180 LIYANRLSGVPSRFTTSSGSGPDFTTISVQAEDLAVYFCQHFRTPTFGSGTKLEI 239  
QY 259 K 259  
DB 240 K 240

RESULT 12  
US-10-071-485-2  
Sequence 2, Application US/10071485  
Publication No. US20030099648A1  
GENERAL INFORMATION:  
APPLICANT: Buysse, Marie-Ange  
Sablou, Erwin  
TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC  
SHOCK  
TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS  
FILE REFERENCE: INNS-015  
CURRENT APPLICATION NUMBER: US/10/071,485  
FILING DATE: 2002-02-07  
PRIOR APPLICATION NUMBER: 09/485,737  
PRIOR FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: PCT/EP 98/05165  
PRIOR FILING DATE: 1998-08-14  
PRIOR APPLICATION NUMBER: EPO 98870139.7  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: EPO 97870122.5  
PRIOR FILING DATE: 1997-08-18  
NUMBER OF SEQ ID NOS: 104  
SOFTWARE: Patentin version 3.0  
SEQ ID NO 2  
LENGTH: 267  
TYPE: PRT  
ORGANISM: Artificial Sequence

FEATURE:  
OTHER INFORMATION: Synthetic  
US-10-071-485-2

Query Match 36.7%; Score 901.5; DB 14; Length 267;  
Best Local Similarity 68.3%; Pred. No. 5.2e-55;  
Matches 172; Conservative 28; Mismatches 49; Indels 3; Gaps 2;

QY 8 LPLMAAKVPKQIOIYVQSGPELKPKPEYKISCKASGYFTNYGMMWVQAQGGGLKMG 67  
DB 12 LLLLAQPMMAQVQLVQSGSELKPKASVYKISCKASGYFTDYGMMWVQAQGGGLKMG 71  
QY 68 WINTSGESTPADPKGRPFSLKTSANTAYLQINNLSKEDMATYFCARMEYHGVYRW 127  
DB 72 WINTYTGSEYVDFKGRFVSLDTSVSAAYLQISSLKADATYFCARGFY--AMDYN 129  
QY 128 GGGTIVYVSSGGGGGGGGGGSDIQLTQSHKFLSTSVGRVSTCKASQDVYNAVAM 187  
DB 130 GGGTIVYVSSGGGGGGGGGGSDIQLTQSPATMASPGRVTLTCSASSI-SYMF 188  
QY 188 YQKKGQSPKLIYASRRYTGVPFRFTSGSGPDFTFTISSVQADLAVYFCQGHFRTP 247  
DB 189 YHQRQGSRRLLIYDTSNLASGVAPAFSGSGSTSYSLTISMEPEDFATYFCOSSSYP 248  
QY 248 FTFGSGTKLEIK 259  
DB 249 FTFGSGTKLEIK 260

## RESULT 13

US-09-766-543-12  
Sequence 12, Application US/09766543  
Patent No. US20020041865A1

GENERAL INFORMATION:  
APPLICANT: Austin, Richard  
APPLICANT: Kwok, Cheuk S.  
APPLICANT: Ring, David B.  
TITLE OF INVENTION: METHODS FOR TREATING TUMORS  
FILE REFERENCE: P01679.002  
CURRENT APPLICATION NUMBER: US/09/766,543  
CURRENT FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: 60/177,258  
PRIOR FILING DATE: 2000-01-20  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: Patent In Ver. 2.0  
SEQ ID NO 12  
LENGTH: 276  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: H520C9sfv plus  
US-09-766-543-12

Query Match 36.7%; Score 897; DB 9; Length 276;  
Best Local Similarity 69.7%; Pred. No. 1.1e-54;  
Matches 168; Conservative 34; Mismatches 35; Indels 4; Gaps 2;

QY 19 QIOLVQSGPELKPKPEYKISCKASGYFTNYGMMWVQAQGGGLKMGWINTSGESTF 78  
DB 23 EIQLVQSGPEVKKPGASVYKISCKASGYTFPANYGMMWVQAQGGGLKMGWINTYGGSTY 82  
QY 79 ADDPKGRPFSLKTSANTAYLQINNLSKEDMATYFCARMEYHGVYRWGGSTYTVVSSG 138  
DB 83 ADDPKGRPFSLKTSANTAYLQINNLSKEDMATYFCARMEYHGVYRWGGSTYTVVSSG 138  
QY 139 GGGSGGGGGGGGGSDIQLTQSHKFLSTSVGRVSTCKASQDVYNAVAMVQAQGGGLKMG 198  
DB 139 GGGSGGGGGGGGGSDIQLTQSHKFLSTSVGRVSTCKASQDVYNAVAMVQAQGGGLKMG 198  
QY 199 LIYASRRYTGVPFRFTSGSGPDFTFTISSVQADLAVYFCQGHFRTPFTFGSGTKLEI 258  
DB 199 LIYASRRYTGVPFRFTSGSGPDFTFTISSVQADLAVYFCQGHFRTPFTFGSGTKLEI 258

QY 259 K 259  
DB 259 K 259

## RESULT 14

US-10-071-485-85  
Sequence 85, Application US/10071485  
Publication No. US20030099648A1

GENERAL INFORMATION:  
APPLICANT: Buyse, Marie-Ange  
APPLICANT: Sablon, Erwin  
TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC  
TITLE OF INVENTION: SHOCK,  
TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS  
FILE REFERENCE: INNS:015  
CURRENT APPLICATION NUMBER: US/10/071,485  
CURRENT FILING DATE: 2002-02-07  
PRIOR APPLICATION NUMBER: 09/485,737  
PRIOR FILING DATE: 2000-02-14  
PRIOR APPLICATION NUMBER: PCT/EP 98/05165  
PRIOR FILING DATE: 1998-08-14  
PRIOR APPLICATION NUMBER: EPO 98870139.7  
PRIOR FILING DATE: 1998-06-18  
PRIOR APPLICATION NUMBER: EPO 97870122.5  
PRIOR FILING DATE: 1997-08-18  
NUMBER OF SEQ ID NOS: 104  
SOFTWARE: Patent In version 3.0  
SEQ ID NO 85  
LENGTH: 541  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: SYNTHETIC  
US-10-071-485-85

Query Match 36.4%; Score 889.5; DB 14; Length 541;  
Best Local Similarity 67.7%; Pred. No. 7.9e-54;  
Matches 168; Conservative 30; Mismatches 47; Indels 3; Gaps 2;

QY 19 QIOLVQSGPELKPKPEYKISCKASGYFTNYGMMWVQAQGGGLKMGWINTSGESTF 78  
DB 1 QIOLVQSGPELKPKASVYKISCKASGYFTDYGMMWVQAQGGGLKMGWINTYGGSTY 60  
QY 79 ADDPKGRPFSLKTSANTAYLQINNLSKEDMATYFCARMEYHGVYRWGGSTYTVVSSG 138  
DB 61 VDDPKGRPFSLKTSANTAYLQINNLSKEDMATYFCARMEYHGVYRWGGSTYTVVSSG 118  
QY 139 GGGSGGGGGGGGGSDIQLTQSHKFLSTSVGRVSTCKASQDVYNAVAMVQAQGGGLKMG 198  
DB 119 GGGSGGGGGGGGGSDIQLTQSPATMASPGRVTLTCSASSI-SYMFYHQRQGSRRLLIYDTS 177  
QY 199 LIYASRRYTGVPFRFTSGSGPDFTFTISSVQADLAVYFCQGHFRTPFTFGSGTKLEI 258  
DB 178 LIYASRRYTGVPFRFTSGSGPDFTFTISSVQADLAVYFCQGHFRTPFTFGSGTKLEI 237  
QY 259 KALEISNS 266  
DB 238 KRTPLGDT 245

## RESULT 15

US-10-071-485-90  
Sequence 90, Application US/10071485  
Publication No. US20030099648A1

GENERAL INFORMATION:  
APPLICANT: Buyse, Marie-Ange  
APPLICANT: Sablon, Erwin  
TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC  
TITLE OF INVENTION: SHOCK,  
TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS  
FILE REFERENCE: INNS:015

```

; CURRENT APPLICATION NUMBER: US/10/071,485
; CURRENT FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 09/485,737
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 90
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
; US-10-071-485-90

```

```

Query Match      36.4%; Score 888.5; DB 14; Length 711;
Best Local Similarity 69.7%; Pred. No. 1.3e-53;
Matches 168; Conservative 27; Mismatches 43; Indels 3; Gaps 2;

```

```

QY 19 QIQVQSGPELKKPGETVYKISCKASGYPTNYGMNWKQAPGQGLKMMGWINTSTGESTF 78
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 473 QVQVQSGSELKKPGASVYKISCKASGYPTDYGMMWVKQAPGQGLKMMGWINTYTGESTY 532
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 79 ADDPKGRFDFSLSTANTAYLQINNKKSDMATYFCARMEVYHGVVPYWGQGTITVSSG 138
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 533 VDDFKGRVFLSDIVSANAYLQISSLKABDTATYFCARRGFY--ANDYWGQGTITVSSG 590
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 139 GGGSGGGSGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDPVYNVAVMYOQKPGQSPKL 198
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 591 GGGSGGGSGGGSDIVLTQSPATWSASPERVTLTCSASSI--SYMFWMYHQRPGQSPRL 649
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY 199 LIYSASSRYTVPSKFTSGSGSPDFTTISVQAEIDLAVYFCQCHFRTPFTFGSGTKLEI 258
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 650 LIYDTSNLASGVPAFSGSGSTSYLTISRMEPDAFYFCHQSSSYPTFGQGTKLEI 709
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
QY .259 K 259
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||
DB 710 K 710
   |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||  |||||

```

Search completed: April 20, 2005, 07:10:38  
 Job time : 61.1549 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 20, 2005, 10:01:29 ; Search time 6166.38 Seconds

(without alignments)  
3622.520 Million cell updates/sec

Title: US-09-596-774-6

Perfect score: 2441  
Sequence: 1 MAMWVTLFLMAAAKVPKQI.....ISTATKDTYDALHMQLAPR 461

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2.1/USPTO\_epool\_h/US09596774/runac\_20042005\_075442\_12508/app\_query.fasta\_1.1230  
-DB=GenEmbl -QFMT=faeap -SUFFIX=ap20.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human4.0.cdi -LIST=45  
-DOCLINK=200 -THR SCORE=pcr -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=ptco -NORM=ext -HEPRTSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09596774 @CGN\_1\_1.663 @runac\_20042005\_075442\_12508 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESQUERY -NEG\_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl:  
1: gb\_da:  
2: gb\_hcg:  
3: gb\_in:  
4: gb\_om:  
5: gb\_ov:  
6: gb\_pat:  
7: gb\_ph:  
8: gb\_pl:  
9: gb\_pr:  
10: gb\_ro:  
11: gb\_rs:  
12: gb\_sy:  
13: gb\_un:  
14: gb\_vl:

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2441	100.0	1479	6	A47281 Sequence 5
2	1506.5	61.7	3058	6	A93936 Sequence 1
3	1503.5	61.6	1284	6	A93938 Sequence 3
4	1315.5	53.9	2012	6	A22539 M.musculus/

5	1315.5	53.9	2012	6	128526 Sequence 13
6	1312.5	53.8	2070	6	A57337 Sequence 7
7	1312.5	53.8	2730	6	A57353 Sequence 23
8	1308	53.6	894	12	XXU64991
9	1308	53.6	1862	6	A50995 Sequence 36
10	1308	53.6	1862	6	AR268367 Sequence
11	1308	53.6	1919	6	A50993 Sequence 34
12	1308	53.6	1919	6	AR268366 Sequence
13	1308	53.6	2214	6	A57341 Sequence 11
14	1308	53.6	2793	6	A57343 Sequence 13
15	1308	53.6	2793	6	A57351 Sequence 21
16	1308	53.6	3177	6	A57347 Sequence 17
17	1305	53.5	1692	6	A50960 Sequence 1
18	1305	53.5	1692	6	AR268336 Sequence
19	1303	53.4	2233	6	A22470 M.musculus/
20	1303	53.4	2233	6	128521 Sequence 6
21	1302	53.3	2718	6	A57333 Sequence 3
22	1302	53.3	2718	6	A57349 Sequence 19
23	1302	53.3	2793	6	A57345 Sequence 15
24	1285.5	52.7	748	6	A22469 M.musculus
25	1285.5	52.7	748	6	A47277 Sequence 1
26	1285.5	52.7	748	6	128520 Sequence 4
27	1013	41.5	726	6	AX100174 Sequence
28	1011.5	41.4	1422	6	A63768 Sequence 32
29	1011.5	41.4	1545	6	A63770 Sequence 34
30	994	40.7	771	6	AX057978 Sequence
31	981.5	40.2	909	6	AR007979 Sequence
32	981.5	40.2	909	6	AR058994 Sequence
33	981.5	40.2	909	6	123444 Sequence 1
34	981.5	40.2	909	6	143364 Sequence 7
35	981.5	40.2	909	6	161419 Sequence 7
36	981.5	40.2	909	6	196033 Sequence 7
37	978.5	40.0	987	6	AX935453 Sequence
38	975.5	39.5	987	6	AX935402 Sequence
39	963.5	39.5	987	6	AX935455 Sequence
40	962.5	39.4	1956	6	A63774 Sequence 38
41	962.5	39.4	2079	6	A63776 Sequence 40
42	960.5	39.3	777	6	AX935404 Sequence
43	942	38.6	1632	12	AY452127 Synthetic
44	942	38.6	1761	12	AY452134 Synthetic
45	942	38.6	3195	12	AY452128 Synthetic

#### ALIGNMENTS

RESULT 1  
A47281 LOCUS A47281 1479 bp DNA linear PAT 07-MAR-1997  
DEFINITION Sequence 5 from Patent WO9530014.  
ACCESSION A47281  
VERSION A47281.1 GI:2301300  
KEYWORDS  
SOURCE  
ORGANISM  
unidentified  
unidentified  
unclassified.

REFERENCE  
1 (bases 1 to 1479)  
Groner, B. and Moritz, D.

AUTHORS  
BIFUNCTIONAL PROTEIN, PREPARATION AND USE  
Patent: WO 9530014-A 5 09-NOV-1995;  
JOURNAL

COMMENT  
CIBA GEIGY AG (CH)  
Other publication ZA 9503440 951102  
Other publication AU 2446995 951129.

#### FEATURES

source

1..1479  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
40..1425  
/codon\_start=1  
/product="LEADER-SCPV (FRP5) :LYT-2 HINGE:ZETA"  
/protein\_id="CA03050.1"  
/db\_xref="GI:2301301"  
/translation="MAMWVTLFLMAAAKVPKQIOLVOSGPBLKKPGETVKSCKASG"

YPTNYGNMVKQAPQGLKMMGINTSGESTPADDEKFRPFSLETSSANAYLOIN  
 NKSDMATYPCARMEVHYGVRYMGCGCTTVYSSGGSSGGSSGGSSGGSDLOLQSH  
 KLSSTSVGRVSIITKASQDVYNAVNAVMTQKFGSGKLEITASSRTGVRPFTGSG  
 SGPDEFITSSVQADLAVYFCQOHRPTFGSGKLEIKALEISNSMYSPVVP  
 LQKVSSTTKPVLRTSPVHPPTGSPORBEDCPRGVYKGLDFLEBPKCYLLDG  
 ILFIYGVITLALYLAKFSRSAETAMNODPNQLYNEINLGRREYDVEKKRADPE  
 MGKQORRNPOEGVYNALQKDMABAYSEICTKGBRRRKGHDGLYQGLSATYMDTY  
 DALHMQTLAPR"

## ORIGIN

Alignment Scores:  
 Pred. No.: 1.18e-180 Length: 1479  
 Score: 2441.00 Matches: 461  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-596-774-6 (1-461) X A47281 (1-1479)

QY 1 MetalaTriPValTriPThreuleuphelaMetalaAlaAlaValProlyGlnile 20  
 DB 40 ATGGCTTGGGTGACCTTCTTCTGATGGACGCTGCCAAGTGCCCAAGAGATC 99  
 QY 21 GlnleuValGlnSerGlyProGlnleuYslyProGlyGlnThrValIleSerCys 40  
 DB 100 CAGTTGGTCACTGAGACCTGAGCTGAAGAGCCCTGGAGAGACAGTCAAGTCTCTGC 159  
 QY 41 LysAlaSerGlyTyrProPheThrSerTyrGlyMetAsnTriPValIleGlnAlaProGly 60  
 DB 160 AAGGCTCTGGGTATCTTTCACAACTATGAAAGAACTGGGTCAACAGGCTCCACAGA 219  
 QY 61 GlnGlyLeuYsTriPMeGlyTyrIleAsnThrSerThriGlySerThrPheAlaAsp 80  
 DB 220 CAGGCTTAAAGTGAATGGGCTGAGATTAACACTCACTGAGAGATCAACATTTCTGAT 279  
 QY 81 AspPheYsGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyrLeuGln 100  
 DB 280 GACTTCAGAGGAGCGGTGTGACTTCTTGTGAAACCTCGCCAAACACTGCTATTTGAG 339  
 QY 101 IleAsnleuLeuYsSerGlnAspMetAlaThrTyrPheCysAlaArgTriPValIleTyr 120  
 DB 340 ATCAACAACCTCAAAAGTAAGACATGCTACATATTTCTGTGCAAGAGGAGGATTTAC 399  
 QY 121 HisGlyTyrValProTyrTriPValGlnGlyThrThrValThrValSerSerGlyGlyGly 140  
 DB 400 CACGGCTACGTTCTTCTTACTGGGGCCCAAGGACACGCTTCTTCTCTGGGGTGGC 459  
 QY 141 GlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGlnSerHis 160  
 DB 460 GGTCTGTGGCGGTGGCTCGGGCGGTCTGTGACATCCAGCTGACCCAGTCTCAC 519  
 QY 161 LysPheLeuSerThrSerValGlyAspArgValSerIleThrCysIlyAspAlaSerGlnAsp 180  
 DB 520 AAATTTCTCTGCTCACTTCACTGAGAGACAGGCTCAGATCACTGCAAGGCCAGTCAGAT 579  
 QY 181 ValIlyAsnAlaValAlaTriPValGlnGlnIlyProGlyGlnSerProIlyLeuLeuIle 200  
 DB 580 GTGATATATGCTGTGCTGCTGATCAACAGAAACAGAGACATCTCTAACTCTGAT 639  
 QY 201 TyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGly 220  
 DB 640 TACTGGGATCTCCCGGTACACAGAGTCCCTTCTCTCTTCACTGCGAGCTCTGGG 699  
 QY 221 ProAspPheThrPheThrIleSerSerValGlnIleGlnAspLeuAlaValIlyPheCys 240  
 DB 700 CCGGATTTCACTTTCACATCAGCAGTGTGAGGCTGAAGCTGGCAGTTATTTCTGT 759  
 QY 241 GlnGlnIleAspPheArgThrProPheThrPheGlySerGlyThrIlyLeuGlnIleValAla 260  
 DB 760 CAGCAACATTTTCTTACTCATTACAGTCTGGGTCAGGAGCAAAATGAGATCAAGCT 819  
 QY 261 LeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGlnIlyVal 280

DB 820 CTAGAGATCAGCAACTCGGTATGACTTCACTTGTCTGTGCGCAAGTCTTTCAGAAAGTG 879  
 QY 281 AsnSerThrThrThrIlyProValIleuArgThrProSerProValIleProthGlyThr 300  
 DB 880 AACTTCACTTACCAAGCCAGTCTGGAACCTCCTCAGTGTGACACCTACCGGAGCA 939  
 QY 301 SerGlnProGlnArgProGlnAspCysArgProArgIlySerValIlyGlyIlyLeu 320  
 DB 940 TCTCAGCCCGCAGAGACCAAGAAATGTCGCGCCGTGCTCAGTGAAGGGAGCCGAGTTG 999  
 QY 321 AspPheLeuGlnAspProIlyLeuCysTyrLeuLeuAspGlyIleLeuPheIleTyrGly 340  
 DB 1000 GACTTTCAGAGATCCCAAACTCTGCTACTTCTGATGGAATCTTCTTATCTTACGA 1059  
 QY 341 ValIleIleThrAlaLeuTyrLeuArgAlaIlyPheSerArgSerAlaGlnThrAlaAla 360  
 DB 1060 GTCATTCATCAGCCCTGTACTGAGAGCAAAATTCAGCAGAGTGCAGAGACTGTGCC 1119  
 QY 361 AsnLeuGlnAspProAsnGlnLeuTyrArgGlnLeuAsnLeuGlyIlyArgArgGlnIlyTyr 380  
 DB 1120 AACTCGAGAGCCCAACAGCTTCAATGAGCTCAATCTAGGGCGAGAGAGAAATAT 1179  
 QY 381 AspValLeuGlnIlyIlyValIlyTyrAsnAlaLeuGlnIlyAspIlySerMetAlaIlyTyr 420  
 DB 1180 GAGCTCTTGGAGAAAGAGCGGGCTCGGATCCAGAGATGGAGAGCAACAGCAGAGAGC 1239  
 QY 401 ArgAsnProGlnGlnIlyValIlyTyrAsnAlaLeuGlnIlyAspIlySerMetAlaIlyTyr 420  
 DB 1240 AGGAACCCCGCAGAGAGCGTATCAATGACCTGACAGAAACAAATGCGAGAGCCTAC 1299  
 QY 421 SerGlnIleGlyThrIlyIlyGlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIlyIly 440  
 DB 1300 AGTCAAGTCCGACAAAGCGAGAGCGAGAGCGAGAGCGAGAGCGAGAGCGAGAGCGAG 1359  
 QY 441 GlyLeuSerThrAlaThrIlyAspThrTyrAspAlaLeuIlyIlyIlyIlyIlyIlyIlyIly 460  
 DB 1360 GGTCTCAGACCTGCAACAGAGACACTTGTGATGAGCCCTGATGAGACCTCGGCCCT 1419  
 QY 461 Arg 461  
 DB 1420 CGC 1422

RESULT 2  
 A93936 3058 bp DNA linear PAT 22-JAN-2000  
 LOCUS  
 DEFINITION Sequence 1 from Patent WO9720938.  
 ACCESSION A93936  
 VERSION A93936.1 GI:6742038  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 REFERENCE 1 (bases 1 to 3058)  
 AUTHORS Rosenthal, F. and Kulmburg, P.  
 TITLE CELLS WITH HYBRID RECEPTOR AND GENE CONSTRUCT WHICH CAN BE  
 CONTROLLED BY SAID HYBRID RECEPTOR, AND USE OF SAID CELLS IN GENE  
 THERAPY  
 JOURNAL Patent: WO 9720938-A, 1, 12-JUN-1997.  
 FEATUERS location/Qualifiers  
 source 1..3058  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

ORIGIN

Alignment Scores:  
 Pred. No.: 1.76e-107 Length: 3058  
 Score: 1506.50 Matches: 299  
 Percent Similarity: 75.27% Conservative: 45  
 Best Local Similarity: 65.43% Mismatches: 86  
 Query Match: 61.72% Indels: 27



```

QY 96 ThrAlaTyrLeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAla 115
DB 244 AACGCTTACATGAGCTGACAGCTGACATCTGAGACTCTGGCGTCTATTATTGTGCA 303
QY 116 ArgTyrGluValTyr---HisGlyTyrValProTyrTyrGlnGlnGlnTyrThrValThr 134
DB 304 AGATACGATTACTACGGGTAGTAGTACTTGTGACTAGTGGGCGCAAGGACCGAGCTCAC 363
QY 135 ValSerSerGlyGlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerAspIle 154
DB 364 GTCTCTCAGGTGAGGCGGTTCAGCGGAGGTGCTGCGGCGGTGCGGATCTCAGGCT 423
QY 155 GlnLeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThr 174
DB 424 GTTGGGACACAG---GAATCTGCATCTACACATCACTGCGTGAACAGTCACTCACT 480
QY 125 CysLysAlaSerGlnAspValTyr-----AsnAlaValAlaTyrTyrGlnGlnLys 191
DB 481 TGTGCTCAAGTACGGGCTGTACAACTACTACTACTGCACTGCGGTCCAGAGAAAA 540
QY 192 ProGlyGlnSerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThrGlyValPro 211
DB 541 CCAGATCATTTATTACATGCTGCTATAGTGTGTACCAACACCGAGCTCCAGGTGCTCT 600
QY 212 SerArgPheThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGln 231
DB 601 GCCAATTTCTCAGGCTCTCCATGTTGAGACAGAGCTGCCCTCACCATCAACAGGGCACAG 660
QY 232 AlaGluAspLeuAlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGly 251
DB 661 ACTGAGGATGAGGCGCAATATATTTCTGCTCATGTGTACAGCAACCATTTGGGTTCGCT 720
QY 252 SerGlyThrLysLeuGlnIleLysAlaLeuGlnIleSerAsnSerValMetTyrPheSer 271
DB 721 GGAGAACCAAACTACTGCTC-----CTCGAG----- 747
QY 272 SerValValProValLeuGlnLysValAsnSerThrThrThrLysProValLeuArgThr 291
DB 748 -----GAATTCATCTACTCAACAGCCGCTCTCGAAGCT 780
QY 292 ProSerProValHisProThrGlyThrSerGlnProGlnAspProGlyAspPro 311
DB 781 CCCCTACCTGTGACGCCCTACCGGACATCTCAGCCCCAGACCAAGAGATTTGCGGCC 840
QY 312 ArgGlySerValLysGlyTyrGlyLeuAspPheLeuGluAspProLysLeuCysTyrLeu 331
DB 841 CTGGCTCAGTGAAGGGAGCCGAGTTGACCTTC-----GACCCCAAACTGTGCTACTTG 894
QY 332 LeuAspGlyIleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLys 351
DB 895 CTAGATGGAACTCTCTTCTATCGAGATCATCATCAAGCCCTGTACTGAGCGAAAA 954
QY 352 PheSerArgSerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlu 371
DB 955 TTCAGCAGAGATGACAGACTGCTGCCAACTGACAGACCCCAACAGCTCTTCAAAAGAG 1014
QY 372 LeuAsnLeuGlyArgArgGlnGluTyrAspValLeuGlnLysLysAspAlaArgAspPro 391
DB 1015 CTCAATCTAGGCGCAAGAGAGAAATATGACTCTTGGAGAGAAAGCGGCTCGGAGATCCA 1074
QY 392 GluMetGlyGlyLysGlnGlnArgArgAsnProGlnGlnGlyValTyrAsnAlaLeu 411
DB 1075 GAGATGGGAGGCAACAGCAGAGAGAGAGAACCCCGAGAAAGCATATACATGACACTG 1134
QY 412 GlnLysAspLysMetAlaGluAlaTyrSerGluIleGlyThrLysGlyLysGlnArgArg 431
DB 1135 CAGAAAGCAACATGACAGAAAGCTTACAGTGAATGCGCAAAAAGCGAGAGCGAGAGA 1194
QY 432 GlyLysGlyHisAspGlyLysLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAsp 451
DB 1195 GGCAAGGGGACGATGGCGCTTTACAGAGGTCTCAGCACTGCCACCAAGACACTTATGAT 1254
QY 452 AlaLeuHisMetGlnThrLeuAlaProArg 461

```

```

DB 1255 GCCCTGATATGACAGACCTGGCCCTCTGC 1284
RESULT 4
LOCUS A22539 2012 bp DNA linear PAT 15-NOV-1994
DEFINITION M.musculus/P. aeruginosa Fv(FRP5)-ETA fusion protein.
ACCESSION A22539
VERSION A22539.1 GI:641549
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences; artificial sequences.
REFERENCE
1 (bases 1 to 2012)
Wels,W.S., Hynes,N.E., Harwerth,I.M., Groner,B., Hardman,N. and
AUTHORS Zwickl,M.
TITLE Recombinant antibodies specific for a growth factor receptor
JOURNAL CIBA-GEIGY AG
FEATURES
location/Qualifiers
source 1..2012
/organism="synthetic construct"
/mol_type="unnasigned DNA"
/db_xref="taxon:32630"
1..1911
/product="Fv heavy and light chain variable domain and
exotoxin A fusion protein"
ORIGIN
Alignment Scores:
Pred. No.: 7,79e-93 Length: 2012
Score: 1315.50 Matches: 280
Percent Similarity: 68.34% Conservative: 20
Best Local Similarity: 63.78% Mismatches: 45
Query Match: 53.89% Indels: 95
DB: Gaps: 10
US-09-596-774-6 (1-461) x A22539 (1-2012)
QY 15 LysValProLysGlnIleGlnLeuValGlnSerGlyProGlnLeuLysProGlyGlu 34
DB 85 AACGTGCTTCTCAGGATCAACACTGACAGCTGCTGACCTGAATGAAGACCTGGAGAG 144
QY 35 ThrValLysIleSerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyr 54
DB 145 ACGTCAAGATCTCTCGCAAGCGCTCTGAGTATCTTCAAACTATGGAATGAATCGG 204
QY 55 ValLysGlnAlaProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGly 74
DB 205 GTGAGAGAGCTCCAGAGACAGGTTTAAAGTGATGGGCTGATTAACACCTCAGTGA 264
QY 75 GluSerThrPheAlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAla 94
DB 265 GAGTCAACATTGCTGATGACTTCAAGGACAGGTTTGACTTCTCTTGGAAACCTTGCC 324
QY 95 AsnThrAlaTyrLeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCys 114
DB 325 AACACTGCTTATTGCAATCAACAACTCAAAAGTGAAGACATGGCTCAATATTTCTGT 384
QY 115 AlaArgTyrGluValTyrHisGlyTyrValProTyrTyrGlnGlnGlnTyrThrValThr 134
DB 385 GCAAGATGGGAGGTTTACACCGGCTACGTTCCTTACTGGGGCGCAAGGACACCGGTAC 444
QY 135 ValSerSerGlyGlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerAspIle 154
DB 445 GTTCTCTGCGGCGGTGCGGTCTGCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTG 504
QY 155 GlnLeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThr 174
DB 505 CAGCTAACCAAGTCTACAAATTTCTGCTCACTTCACTGAGGAGACAGGATCGCATCAC 564
QY 175 CysLysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGln 194

```

```

Db      565 TCAGAGCCAGTCAGGATGTATATGCTGTGCTGCTGATCAACAGAAACGAGACAA 624
Oy      195 SerProlybLeuLeuIleTyrSerAlaSerArgTyrThrGlyValProSerArgPhe 214
Db      625 TCTCCTAAACTTCTGATTTACTCGGCATCTCCCGGTACACGTGAGTCCCTTCTGCTTC 684
Oy      215 ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnaGluAsp 234
Db      685 ACTGCAGATGCTCTGGGCGGATTTCACTTCCATCAGCATGAGTGTGACGAGCTGAAGAC 744
Oy      225 LeuAlaValTyrPheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThr 254
Db      745 CTGGCAGATTATTTCTGTACAGCAACATTTTCGTACTCCATTCAGCTTCGGCTCGGGAGCA 804
Oy      255 LysLeuGlnIleLeuAlaLeuGlu-----IleSerAsnSerValMetTyrPheSer 271
Db      805 AAATTGGAGATCAAAAGCTCTAGAGCGGCGGACCTGGCGGCTGACCGGACGACGAGGC 864
Oy      272 SerValValProValLeuGlnIleValAsnSerThrThrLys----- 286
Db      865 TGCCACCTCCGCGT-----GAGACTTTCACCCGTCATCGCCAGCCGCGC 909
Oy      287 -----ProValLeuArgThrProSerProValHis 296
Db      910 GGCTGGGAACAACGTGACAGTGGCGGTATCCGGTGCACGCGCTGGT-CGCGCTCTA--- 965
Oy      297 ProThrGlyThrSer-----GlnProGlnArgProGluAspCysArgProArgGlySer 314
Db      966 CCTGGCGGGCGACGTGCATGAGACCGAGTCCAGCAGGTGATCCGCAACGCCCTGGCCAG 1025
Oy      315 ValIysGlyThrGlyLeuAspPheLeuGluAspProLysLeuCysEtyrLeuLeuAspGly 334
Db      1026 CCCCCGACGCGGCGGCGGACCTGGCGGACGATCCG----- 1061
Oy      335 IleuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArg 354
Db      1062 -----CGA 1064
Oy      355 SerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeu 374
Db      1065 GCAGCGGAGCAGCGCCGCTGCGCCCTGACCT----- 1097
Oy      375 GlyArgTargGluGluTyrAspValLeuGluLysArgAlaArgAspProGluMetGly 394
Db      1098 GGGCGCCCGCCGAG-----CGAGCCCTTCGCTCG-----GCA 1130
Oy      395 GlyLysGlnGlnArgArgArgAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAsp 414
Db      1131 GGGCAGCGGCAACGACGA----- 1148
Oy      415 LysMetAlaGlnAlaTyrSerGluIleGlyThrLysGlyGluArgArgGlyLys 433
Db      1149 -----GGCGCGCGGCGCAACGCGCAGCTGTGTAG 1178

RESULT 5
LOCUS      128526      2012 bp      DNA      linear      PAT 06-FEB-1997
DEFINITION Sequence 13 from patent US 5571894.
ACCESSION 128526
VERSION    128526.1 GI:1819302
KEYWORDS
SOURCE
ORGANISM   Unknown.
            Unclassified.
REFERENCE 1 (bases 1 to 2012)
AUTHORS   Wells,W.S., Hynes,N.E., Harweth,I.-M., Groner,B., Hardman,N. and
            Zwickl,M.
            Recombinant antibodies specific for a growth factor receptor
            Patent: US 5571894-A 13 05-NOV-1996;
            Location/Qualifiers
            1..2012
            /organism="Unknown"
            /mol_type="unassigned DNA"

```

```

ORIGIN
Alignment Scores:
Pred. No.:      7,796-93
Score:          1315.50
Percent Similarity: 68.34%
Best Local Similarity: 63.78%
Query Match:    53.89%
DB:             6
               Gaps: 10

US-09-596-774-6 (1-461) x 128526 (1-2012)
Oy      15 LysValProLysGlnIleGlnLeuValGlnSerGlyProGluLeuLysAspProGlyGlu 34
Db      85 AAGCTACGCTTCAGGTCAACTGACAGTGTGACCTGAAGTGAAGAGCTCGGAGAG 144
Oy      35 ThrValLysIleSerCysAlaLysSerGlyTyrProPheThrAsnTyrGlyMetAsnTrp 54
Db      145 ACAGTCAAGATCTCTGCAAGGCGCTCGGTATCTCTTACAACTATGATGAATCTGG 204
Oy      55 ValIysGlnAlaProGlyGlnGlyLeuLysTrpMetGlyTrpIleAsnThrSerThrGly 74
Db      205 GTGAAGCAGGCTCCAGACAGGCTTTAAAGTGAAGGCTGATTAACACCTTCACCTGSA 264
Oy      75 GluSerThrPheAlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAla 94
Db      265 GAGTCAACATTTGCTGATGATCTTCAAGGAGCGTTTGACTTCTCTTGGAAACCTCTGCC 324
Oy      95 AsnThrAlaTyrLeuGlnIleAsnAlaLeuLysSerGluAspMetAlaThrTyrPheCys 114
Db      325 AACACTGCTATTGTCAGATCAACAACTCAAAAGTGAAGCATGCTATACATATTTCTGT 384
Oy      115 AlaArgTrpGluValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThr 134
Db      385 GCAGATGAGAGGTTTACACGCGCTTACGTTCTTACTGAGGCAAGGACACCGGTCAC 444
Oy      445 GTTTCCTCTGCGCGGCGGCGGTTCTGTGTGGCGTGGCTCGGCGGCGGTTCTGACATC 504
Oy      155 GlnLeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThr 174
Db      505 CAGCTGACCCAGCTCTCAAAATTCCTGTTCACCTTCAAGTGAAGACAGGCTGACATCAC 564
Oy      175 CysLysAlaSerGlnAspValTyrAsnAlaValAlaTyrPyrGlnGlnLysProGlyGln 194
Db      565 TCAGAGCCAGTCAGAGATGTATATGCTGTGCTGCTGATCAACAGAAACGAGACAA 624
Oy      195 SerProlybLeuLeuIleTyrSerAlaSerArgTyrThrGlyValProSerArgPhe 214
Db      625 TCTCCTAAACTTCTGATTTACTCGGCATCTCCGGTACACGTGAGTCCCTTCTGCTTC 684
Oy      215 ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnaGluAsp 234
Db      685 ACTGCAGATGCTCTGGGCGGATTTCACTTCAACATCAGACAGTGTGACGAGCTGAAGAC 744
Oy      225 LeuAlaValTyrPheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThr 254
Db      745 CTGGCAGATTATTTCTGTACAGCAACATTTTCGTACTCCATTCACCTTGGCTCGGGAGCA 804
Oy      255 LysLeuGlnIleLeuAlaLeuGlu-----IleSerAsnSerValMetTyrPheSer 271
Db      805 AAATTGGAGATCAAAAGCTCTAGAGCGGCGGACCTGGCGGCTGACCGGACGACGAGGC 864
Oy      272 SerValValProValLeuGlnIleValAsnSerThrThrLys----- 286
Db      865 TGCCACCTCCGCGT-----GAGACTTTCACCCGTCATCGCCAGCCGCGC 909
Oy      287 -----ProValLeuArgThrProSerProValHis 296
Db      910 GGCTGGGAACAACGTGACAGTGGCGGTATCCGGTGCACGCGCTGGT-CGCGCTCTA--- 965
Oy      297 ProThrGlyThrSer-----GlnProGlnArgProGluAspCysArgProArgGlySer 314

```

```

Db      966 CCTGGCGCGCGCGCTGTATGAAACGAGTCGACGACGATGCCAAGCCCTGGCCAG 1025
QY      315 ValLysGlyThrGlyLeuAspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGly 334
Db      1026 CCCCAGCGCGCGCGCGACCTGGCGCAGACGATCCG----- 1061
QY      335 IleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArg 354
Db      1062 -----CGA 1064
QY      355 SerAlaGluThrAlaAlaAsnLeuGluAspProAsnGlnLeuTyrAsnGlnLeuAsnLeu 374
Db      1065 GAGCGCGGACGACGCCCTGTGCGCTGACCT----- 1097
QY      375 GlyArgArgGluGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGly 394
Db      1098 GCGCGCGCGCGAGC-----CGAGCGCTTCGTCCG-----GCA 1130
QY      395 GlyLysGlnGlnArgArgAspProGlnGluGlyValTyrAsnAlaLeuGlnLysAsp 414
Db      1131 GGGCAGCGCGCAACGACGA----- 1148
QY      415 LysMetAlaGluAlaTyrSerGluIleGlyThrLysGlyLysArgArgGlyLys 433
Db      1149 -----GGCGCGCGCGCGCAACGCGCGACGTCGTGAG 1178

RESULT 6
A57337 2070 bp DNA linear PAT 03-MAR-1998
LOCUS Definition 7 from Patent EP0739984.
ACCESSION A57337
VERSION A57337.1 GI:3713216
KEYWORDS
SOURCE synthetic construct
ORGANISM other sequences; artificial sequences.
REFERENCE
1. Weis, W.D., Schmidt, M. and Groner, B.D.
   Bivalent polypeptides containing at least two domains
   Patent: EP 0739984-A 7 30-OCT-1996;
   SAN TUMORFORSCHUNGS GMBH (DE)
FEATURES
    source          location/qualifiers
    1..2070
    /organism="synthetic construct"
    /mol_type="unassigned DNA"
    /db_xref="taxon:32630"
    /clone="PSM202-5"
    1..1956
    /note="unnamed protein product"
    /codon_start=1
    /transl_table=1
    /protein_id="CAA03441.1"
    /db_xref="GI:3713217"
    /translation="MKKTATATATAVLAGFATVAQADYKDDDKLHHHHKLOVLOO
    SGEPEKKEGETVTKSCASGYPFTNYGNWVKAPGGGLKMWGINSTSESTFADPF
    KGRDFSLSTANTAYLIQINLKSEDAVYFCARMEYHGVPMVGGTIVTVSSGGG
    GSGGGGSGGSGDIQLOSHKFLSTSDVRYSTCKASQDVYNAVYVQKPGSGPYL
    LIYASSRYTGVPSRFPTGSGGSDPFTTISVQAEADLAVYFCQHPFTPTFGSTGL
    EIKALEHHHHHLEGGSLAALTAAHQAQCHLPLETFRHRRORGRMGLEOCQVPRQVLA
    LYLAARLSMNQDVIVRNALASPGSGDLGAIATREQEQARLAIITLAAASSERVRQG
    TGNDEAGANADVSLTFCVPAAGCAAPADSGDALBERNPTGAEFLDGDGDAVSTR
    GTQWTVERRLQAHRQLBERGYFVGHGTFLENAQSIVFGVARSQDDADAIWRGY
    IAGDPALAYGAQDQEDPARGLRNGALRLVYRPSLPGVYRSLTLTAPEAAGEVE
    RLIGHPLRLDAITGPEEGRLETILGWPLAERTVIVPSAIPTDPRNVGDLDPSS
    IPRKQALISALPDYASQPKPKPRDLK"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 1.38e-92 Length: 2070
Score: 1312.50 Matches: 284
Percent Similarity: 68.59% Conservative: 13
Best Local Similarity: 65.59% Mismatches: 55

```

```

Query Match: 53.77% Indels: 83
DB: 6 Gaps: 8
US-09-596-774-6 (1-461) x A57337 (1-2070)
QY      19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 38
Db      115 CAGGTACAACTGACGACAGTCTGACCTGAAGAAAGAGAGAGAGAGAGAGAGAGAGAGAG 174
QY      39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 58
Db      175 TCTGTCAAGAGCGCTGTGGGTATCTTTCACAAACATAGAAATGAACCTGGGTGAAGCAGCT 234
QY      59 ProGlyGlnGlyLeuLysTrpMetGlyTyrTrpLeuSerThrArgLysGlySerThrPhe 78
Db      235 CCAGACAGAGGTATTAAGTGGATGGCTGGATTAAACCTCCACTGGAGAGTCAACATTT 294.
QY      79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
Db      295 GCTGATGACTTCAGAGGACCGTTGACTTCTCTTGGAAACCTTGCACACACTGCCTAT 354
QY      99 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 118
Db      355 TTGCAGATCAACAACTCAAAAGTGAAGACATGGCTACATATTTCTGTGCAAGATGGAG 414
QY      119 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 138
Db      415 GTTACACAGCGCTACGTTCTTCTGAGGACAGGACACAGGACGTTCTCTGAGC 474
QY      139 GlyLysGlySerGlyGlyGlySerGlyGlyGlySerAlaIleGlnLeuThrGln 158
Db      475 GGTGGCGGTTCTGGTGGCGGTGGCTCCGCGGTTGCGGCTTCAACATCCAGTCGACCCAG 534
QY      159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178
Db      535 TCTCAAAATTCCTGTCCTCACTCAGTAGAGAGAGGCTCAGATCACTGCAAGGCGCAGT 594
QY      179 GlnAspValTyrAsnAlaValAlaTyrGlnGlnLysProGlyLysSerProLysLeu 198
Db      595 CAGAGTGTATATAGTGTGTGCTGCTGTATCAACAGAAACAGAGACATCTCTTAACCT 654
QY      199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db      655 CTGATTTACTCGGCATCTCTCCGCTACCTGAGTCCCTTCTGCTTCACTGGCAGCTGGC 714
QY      219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 238
Db      715 TCTGGCGCGGATTTCACTTTCACCATCAGCATGTCAGGCTGAAGACCTGGGCAATTAT 774
QY      239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 258
Db      775 TTCTGTACAGCAACATTTTGTCTTCACTTCACTTTCAGCTTGGCTCGGGGACAAATTTGAGATC 834
QY      259 LysAlaLeuGluLysSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278
Db      835 AAAGCTCTTAGAGCA----- 848
QY      279 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 298
Db      849 -----CAATCATCAACCATCACTAG 869
QY      299 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGlyThr 318
Db      870 GGG-----CGGACGCTGCGCGCGGTGAC-----CGGACACAGGCTGCGACCTGCGGCT 920
QY      319 GlyLeuAspPheLeuGlnAspProLysLysCysTyrLeuLeuAspGlyIleLeuPheIle 338
Db      921 GGAAGCTTTCACCGCTCATCGCAGCGCGCGGCTGGGAACAACATGACAGTGGCGCTTA 980
QY      339 TyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSer----- 353
Db      981 TCCGATGACAGCGGCTGTGCGCTTCACTGCGCGCGGACATCTGTATGAAACAGAGTCGA 1040

```



```

RESULT 8
XXU64991
LOCUS      894 bp      DNA      linear      SYN 11-FEB-2002
DEFINITION Synthetic construct anti-human erbB-2 receptor bifunctional
            single-chain antibody-phosphatase fusion protein, partial cds.
ACCESSION  U64994
VERSION     U64994.1
KEYWORDS    GI:1545865
SEGMENT     1 of 2
SOURCE      synthetic construct
ORGANISM    other sequences/artificial sequences.
REFERENCE   1 (bases 1 to 894)
AUTHORS     Wels,W., Harwerth,I.M., Zwickl,M., Hardman,N., Groner,B. and
            Hynes,N.E.
TITLE       Construction, bacterial expression and characterization of a
            bifunctional single-chain antibody-phosphatase fusion protein
            targeted to the human erbB-2 receptor
JOURNAL     Biotechnology 10, 1128-1132 (1992)
REMARK      GenBank staff at the National Library of Medicine created this
            entry [NCBI glibseq 115269] from the original journal article.
FEATURES
            location/Qualifiers
            1..894

```

Alignment Scores:	
Pred. No.:	1,08e-92
Score:	108.00
Percent Similarity:	94.328
Best Local Similarity:	93.188
Query Match:	53.588
8:	12
Length:	894
Matches:	246
Conservative:	3
Mismatches:	1
Indels:	12
Gaps:	1

[illegible]

RESULT 9				
LOCUS	A50995			
DEFINITION	A50995	1862 bp	DNA	linear
ACCESSION	Sequence 36 from Patent WO9613599.			
VERSION	A50995			
KEYWORDS	A50995.1	GI:2303794		
SOURCE	unidentified			
ORGANISM	unidentified			
REFERENCE	unclassified.			
AUTHORS	1 (bases 1 to 1862)			
TITLE	Wels, W. and Pominaya, J.			
JOURNAL	NUCLEIC ACID TRANSPER SYSTEM			
	Patent: WO 9613599-A 36 09-MAY-1996;			
	WELS WINFRIED (DE)			
COMMENT	Other publication AU 3926695 960523.			
FEATURES	Location/Qualifiers			



source  
1. .1862  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
CDS  
1. .1854  
/note="unnamed protein product"  
/codon\_start=1  
/protein\_id="CAA03250.1"  
/db\_xref="GI:2303795"

/translation="MDYKDDDDKLLHHHHKLLSSIDQACDLCRLKXKSEKPK  
AKLKNNECRISPKTKRSPLTRALUTEVESLERLEQLFLIPREDIDMLKQDLSL  
QDIALKALGLFQDNDVNDKDAVDRLASVETDPLRLRHRISATSSSESSKQROL  
TVSSSLAVGSSLSICINDMDVIRDKTKTISLKEHGPDKNMSSESPKTSSEKAKO  
LLEEFHOTALEHEPELSELTGTNPVFAAGAYAAVAVAOVIDSEADNLEKTTAA  
LSILPDISVMGILADGAVHNTEETIVASISLISLMAVQALPLVGEIVDIGPAAYNV  
ESIINLFOVHNSTYRPRAYSPGVSDIGDKLOVLOQSGELKPKGTIVKISCASTP  
TNYGNWVQAPQGLKMWGMTINTSGTSTGSDDRGRDSELTSTANTAYLIQNNLK  
SEDMATYFCARWEVYHGVYVYWGQITTVTSVSGSGSGSGSGSDIQLTQSHKFL  
STSVDRVSTICKASQDVAVNAVAYQKRGQSKLLIYSASRYGVSFRFTGSGSGSP  
DPTFTSSVQAEADLAVYFCQGHFRTPFTFGSTKLEIYALDELDSERFSA"

## ORIGIN

Alignment Scores:  
Pred. No.: 2,71e-92 Length: 1862  
Score: 1308.00 Matches: 245  
Percent Similarity: 97.64% Conservative: 3  
Best Local Similarity: 96.46% Mismatches: 6  
Query Match: 53.58% Indels: 0  
DB: 6 Gaps: 0

US-09-596-774-6 (1-461) x AS0095 (1-1862)

Qy 19 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValIle 38  
Db 1090 CAGGTACACAGCTGACAGCTGACCTGAAAGAAAGCTGAGAGACACTCAAGATC 1149  
Qy 39 SerCysValAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58  
Db 1150 TCCTGCAGAGGCTCTGGGTATCTTTCACAACTATGAAATGAACTGGGTGAAGCAGCT 1209  
Qy 59 ProGlyGlnGlyLeuValTyrMetGlyTyrPheAsnThrSerThrGlyIleSerThrPhe 78  
Db 1210 CCAGGACAGGGTTTAAAGTGAAGTGGCTGATTAACCTCCAGCTGAGAGCAATTT 1269  
Qy 79 AlaAspAspPheLeuValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98  
Db 1270 GCTGATGACTTCAGAGGAGCGTTTGAATCTTCTTGGAAACCTCGCCAACTGCTCAT 1329  
Qy 99 LeuGlnIleAsnAsnLeuValSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 118  
Db 1330 TTGCAGATCAACACCTCAAAAGTGAAGATGCTATCTTCTGCGAAGATGGAG 1389  
Qy 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 138  
Db 1390 GTTACACAGGCTACGTTCTTACTGCGGACCAAGGACCAAGCTGCTTCTGCGC 1449  
Qy 139 GlyIleGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158  
Db 1450 GGTGGCGGTTCTGGTGGCGGTGCTCCGCGGTGCGGTTCGACATCCAGCTGACCCAG 1509  
Qy 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaAlaSer 178  
Db 1510 TCTCAAAATTCCTTCCACTTCAGTAGAGACAGGCTCAGATCACTCGCAAGGCCAGT 1569  
Qy 179 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 198  
Db 1570 CAGAGTGTATATATGCTGCTGCTGATCAACAGAAACAGAGCAATCTCTTAACAT 1629  
Qy 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
Db 1630 CTGATTACTCGGCAATCTCCCGGTACACTGGAAGTCCCTTCTGCGTTCACTGGCAGTGGC 1689  
Qy 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238

Db 1690 TCTGGGCGGATTCACCTTACCATCAGACAGTGTGAGAGCTGAGACCTGGCAGATTAT 1749  
Qy 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 258  
Db 1750 TTCTGTACAGCAATTCCTGATCTTCATTCACGTTCTGGGGACAAATTTGAGATC 1809  
Qy 259 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSer 272  
Db 1810 AAAGCTTAGAGATCTCTCAGTAGAGAAAGATTTTACGCC 1851

## ORIGIN

Alignment Scores:  
Pred. No.: 2,71e-92 Length: 1862  
Score: 1308.00 Matches: 245  
Percent Similarity: 97.64% Conservative: 3  
Best Local Similarity: 96.46% Mismatches: 6  
Query Match: 53.58% Indels: 0  
DB: 6 Gaps: 0

US-09-596-774-6 (1-461) x AR268367 (1-1862)

Qy 19 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValIle 38  
Db 1090 CAGGTACACAGCTGACAGCTGACCTGAAAGAAAGCTGAGAGACACTCAAGATC 1149  
Qy 39 SerCysValAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58  
Db 1150 TCCTGCAGAGGCTCTGGGTATCTTTCACAACTATGAAATGAACTGGGTGAAGCAGCT 1209  
Qy 59 ProGlyGlnGlyLeuValTyrMetGlyTyrPheAsnThrSerThrGlyIleSerThrPhe 78  
Db 1210 CCAGGACAGGGTTTAAAGTGAAGTGGCTGATTAACCTCCAGCTGAGAGCAATTT 1269  
Qy 79 AlaAspAspPheLeuValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98  
Db 1270 GCTGATGACTTCAGAGGAGCGTTTGAATCTTCTTGGAAACCTCGCCAACTGCTCAT 1329  
Qy 99 LeuGlnIleAsnAsnLeuValSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 118  
Db 1330 TTGCAGATCAACACCTCAAAAGTGAAGATGCTATCTTCTGCGAAGATGGAG 1389  
Qy 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 138  
Db 1390 GTTACACAGGCTACGTTCTTACTGCGGACCAAGGACCAAGCTGCTTCTGCGC 1449  
Qy 139 GlyIleGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158  
Db 1450 GGTGGCGGTTCTGGTGGCGGTGCTCCGCGGTGCGGTTCGACATCCAGCTGACCCAG 1509  
Qy 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaAlaSer 178  
Db 1510 TCTCAAAATTCCTTCCACTTCAGTAGAGACAGGCTCAGATCACTCGCAAGGCCAGT 1569

QY 179 GlnAspValTyrAsnAlaValAlaTPrIYrGlnGlnLysProGlyGlnSerProLysLeu 198  
DB 1570 CAGGATGTATATATGCTTCTGCTGCTATCAACGAAACAGACAACTCTTAACCTT 1629  
QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
DB 1630 CTGATTTACTCCGCATCTCCCGGTACAGTGCCTCTTCTGCTTCACTGGCAGTGGC 1689  
QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 238  
DB 1690 TCTGGCCGCGATTTCACTTCAACATCAGACAGTGCAGGCTGAAGACCTGGCAGTTTAT 1749  
QY 239 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyTyrIleLeuGluIle 258  
DB 1750 TTCTGTACGACAACTTTTCTGATCTCATTCACGTTTCGCTCGGAGCAAAATTGAGATC 1809  
QY 259 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer 272  
DB 1810 AAGCTTAGAGATCTCTCGAGTGAAGAGATTTCAGCC 1851

## RESULT 11

LOCUS A50993 1919 bp DNA linear PAT 07-MAR-1997  
DEFINITION Sequence 34 from Patent WO9613599.  
ACCESSION A50993  
VERSION A50993.1 GI:2303792  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified.  
REFERENCE 1 (bases 1 to 1919)  
AUTHORS Wels,W. and Fomlnaya,J.  
TITLE NUCLEIC ACID TRANSFER SYSTEM  
JOURNAL Patent: WO 9613599-A 34 09-MAY-1996;  
WELLS WINFRIED (DE)  
COMMENT Other publication AU 3926895 960523.  
FEATURES  
source Location/Qualifiers  
1..1919  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
64..1908  
/note="unamed protein product; Protein sequence is in  
conflict with the conceptual translation"  
/codon\_start=1  
/protein\_id="CAA03249.1"  
/db\_xref="GI:2303793"  
/translation="DYKDDDDKLLHHHHKLLSIEQACDLCRLKTKCKSEKPKCK  
CLKNWBCRYSPKTKRSPVLAHLTEVESRLERLEQLFLIPREDIDMLKMSLD  
IKALLTGLFVQDNNKDAVTDRLASVEMPELTLRHRISATSSSESSNGKRLTV  
SSSLAVGSSLSCLNDMDVIRDKTKIESLKEHGPISSESPNKTVSEKAKOYL  
EEFOTALEHPELSELKTVGTNPVPGANVAAVAVOYIDETADNIEKTPALS  
ILPEIGSMGILADAVHHNTEELVAGSIALSSMAQAALPLGSLVITIGRAVNPVS  
IINFOVYHNSYNRPAYSPGVDGIDKLOVLOQSGPELKKEGTVKISCSAGSPFTN  
YGMNVKQAPQEQGLKMGWINTSTGESIFADDFGRDFSLSTANAYVLIQINLKE  
DMATVFCARMEVHGIVYVWGQCTTVVSSGGSGGGSGGSDILOQSHFLEST  
SVGRVSLITCKASQDVVNAVAYOQKPGSKLIIYASSRYTGVPERFTGSGGPDF  
TFTISSVQAEADLAVYFCQHRPFTFGSGTKLEIKALDELSSRRRSA"

## CDS

1..1919  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
64..1908  
/note="unamed protein product; Protein sequence is in  
conflict with the conceptual translation"  
/codon\_start=1  
/protein\_id="CAA03249.1"  
/db\_xref="GI:2303793"  
/translation="DYKDDDDKLLHHHHKLLSIEQACDLCRLKTKCKSEKPKCK  
CLKNWBCRYSPKTKRSPVLAHLTEVESRLERLEQLFLIPREDIDMLKMSLD  
IKALLTGLFVQDNNKDAVTDRLASVEMPELTLRHRISATSSSESSNGKRLTV  
SSSLAVGSSLSCLNDMDVIRDKTKIESLKEHGPISSESPNKTVSEKAKOYL  
EEFOTALEHPELSELKTVGTNPVPGANVAAVAVOYIDETADNIEKTPALS  
ILPEIGSMGILADAVHHNTEELVAGSIALSSMAQAALPLGSLVITIGRAVNPVS  
IINFOVYHNSYNRPAYSPGVDGIDKLOVLOQSGPELKKEGTVKISCSAGSPFTN  
YGMNVKQAPQEQGLKMGWINTSTGESIFADDFGRDFSLSTANAYVLIQINLKE  
DMATVFCARMEVHGIVYVWGQCTTVVSSGGSGGGSGGSDILOQSHFLEST  
SVGRVSLITCKASQDVVNAVAYOQKPGSKLIIYASSRYTGVPERFTGSGGPDF  
TFTISSVQAEADLAVYFCQHRPFTFGSGTKLEIKALDELSSRRRSA"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,82e-92 Length: 1919  
Score: 1308.00 Matches: 245  
Percent Similarity: 97.64% Conservative: 3  
Best Local Similarity: 96.46% Mismatches: 6  
Query Match: 53.58% Indels: 0  
DB: Gaps: 0

US-09-596-774-6 (1-461) x A50993 (1-1919)

QY 19 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysLysProGlyGlnThrValIle 38  
DB 1147 CAGGTACACTGCAGCACTGTGACCTGAAGAAAGCTGAGAGACAGTCAAGATC 1206

QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTPrValLysGlnAla 58  
DB 1207 TCTGCAAGGCGCTCTGGGTATCTTTTCACAAACATATGAAATGAACCTGGGAGACAGCT 1266  
QY 59 ProGlyGlnGlyLeuLysTPrMetGlyTPrIleAsnThrSerThrGlyGlnSerThrPhe 78  
DB 1267 CCAGACAGGCTTTAAAGTGAATGGCTGATTAACCTTCACTGAGAGATCAACATTT 1326  
QY 79 AlaAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 98  
DB 1327 GCTGATGACTTCAAGGACCGGTTGACTTCTTTGAAACCTCTGCAACACTGCTTAT 1386  
QY 99 LeuGlnIleAsnAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTPrGln 118  
DB 1387 TTGCAGATCAACAACCTCAAAAAGTGAAGACATGCTCAATTTCTGTGAAAGTGGAG 1446  
QY 119 ValTyrHisGlyTyrValProTyrTPrGlyGlnGlyThrThrValThrValSerSerGly 138  
DB 1447 GTTACACACGCTACGTTCTTCTACTGGGSCCAAGGACCAACGTCACCTTCTCTGGC 1506  
QY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 158  
DB 1507 GGTGGCGGTTCTGGTGGCGGCTCCGCGGTGGCGGTTCTGACATCCAGCTGACCCAG 1566  
QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178  
DB 1567 TCTCAAAATTCCTGTCACCTCACTGAGAGACAGGCTCAGATCCTCAAGGCCAGT 1626  
QY 179 GlnAspValTyrAsnAlaValAlaTPrTyrGlnGlnLysProGlyGlnSerProLysLeu 198  
DB 1627 CAGGATGTATATATGCTTGTGCTGGTGCCTGGTGCCTGGTGCCTGGTGCCTGGTGCCT 1686  
QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
DB 1687 CTGATTTACTCCGCATCTCCCGGTACAGTGCCTCTTCTGCTTCACTGGCAGTGGC 1746  
QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 238  
DB 1747 TCTGGCCGCGATTTCACTTCAACATCAGACAGTGCAGGCTGAAGACCTGGCAGTTTAT 1806  
QY 239 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyTyrIleLeuGluIle 258  
DB 1807 TTCTGTACGACAACTTTTCTGATCTCATTCACGTTTCGCTCGGAGCAAAATTGAGATC 1866  
QY 259 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer 272  
DB 1867 AAGCTTAGAGATCTCTCGAGTGAAGAGATTTCAGCC 1908

## RESULT 12

LOCUS AR268366 1919 bp DNA linear PAT 10-APR-2003  
DEFINITION Sequence 34 from patent US 6496233.  
ACCESSION AR268366  
VERSION AR268366.1 GI:29698716  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.

REFERENCE 1 (bases 1 to 1919)  
AUTHORS Wels,W. and Fomlnaya,J.  
TITLE Nucleic acid transfer system  
JOURNAL Patent: US 6496233-A 34 24-DEC-2002;  
FEATURES Location/Qualifiers  
source 1..1919  
/organism="unknown"  
/mol\_type="genomic DNA"

## ORIGIN

## Alignment Scores:

Pred. No.: 2,82e-92 Length: 1919  
Score: 1308.00 Matches: 245  
Percent Similarity: 97.64% Conservative: 3



159 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
655 CTGATTACTCGGCATCTCCCGTACCTGAGTCCCTTCTCGCTTCACTGGCAGTGGC 714  
219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238  
715 TCTGGGCGGATTTCACTTCACTTCACTGAGCAGTGTGAGAGGCTGGAAGACTGGCAGTTAT 774  
229 PheGlyGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrIleLeuGlnIle 258  
775 TCTGTGACCAACATTTTGTGACTCCATTCACGTTCTGGGTGGGACAAATTTGAAGATC 834  
239 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278  
835 AAAGCTCTAGACA----- 848  
279 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 298  
849 -----CCATCATCACCATCACCTAGA 869  
299 GlyThrSerGlnProGlnArgProGlnAspGlySerArgProArgGlySerValIleGlyThr 318  
870 GGG---CGGACGCTGGCGCGCTGAC-----CGCGACACAGGCTGCCACTGGCGCT 920  
319 GlyLeuAspPheLeuGlnAspProLysLeuGlyTyrLeuLeuAspGlyIleLeuPheIle 338  
921 GGAGACTTTCACCCGTCATCCGACCGCGCGGCTGGGACAACTGAGAGCAGTGGCGCTA 980  
339 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 358  
981 TCCGCTGACGCGCGTGTGCTGCTTACCTGCGCGCGGACCTGCA----- 1026  
358 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuValArgArgG1 378  
1027 -----TGAACCAAGTTCGACCA-GG1----- 1046  
378 uGluTyrAspValLeuGlnIleLysArgAlaArgAspProGluMetGlyGlyLysGlnG1 398  
1047 -----GATCGGACAGCCCTGGCCAGCCCG 1072  
398 nArgArgArgAsnPro-GlnGlnGlyValTyrAsnAlaLeuGlnIleLysAspIleAspIle 418  
1073 GCAGCGGCGGACCTGGGCGGACAGATCCCGAGACCCGAGAGGCGCGCTGGCGCC 1132  
418 IuaIaTyrSerGlnIleGlyThrLysGlyGluArgArgG1LysG1YHisAsp 436  
1133 TGACCCCTGGCGCGCGGACGAGCCGCTTCTCGCGAGGCGACCGGCAACGAC 1188

RESULT 14  
A57343 2793 bp DNA linear PAT 03-MAR-1998  
LOCUS Sequence 13 from Patent EP0739984.  
DEFINITION A57343  
ACCESSION A57343  
VERSION A57343.1 GI:3713222  
KEYWORDS  
SOURCE synthetic construct  
ORGANISM synthetic construct  
OTHER SEQUENCES; artificial sequences.

REFERENCE  
1 Wels,W.D., Schmidt,M. and Groner,B.D.  
AUTHORS Bivalent polypeptides containing at least two domains  
JOURNAL Patent: EP 0739984-A 13 30-OCT-1996;  
SAN TUMORFORSCHUNGS GMBH (DE)  
FEATURES  
source location/Qualifiers  
1..2793  
/organism="synthetic construct"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32630"  
/clone="PMS238-5-225"  
1..2679  
/note="unnamed protein product"  
/codon\_start=1  
/transl\_table=11

/protein\_id="CA03444.1"  
/db\_xref="GI:3713223"  
/translation="MKRTALAVLAGPVAQADYKDDDKLHHHHHKLQVQLQO  
SGPELKRGEITVKISCKASGYPTNYGMNWNKQAPGGGLKMKMGINTSTESTPADP  
KGRDFSLSTSAAYLQINNLSKEDATYFCARMEYHYGVYWGQGVTVVSGG  
GGGGGGGGSDILOIQTSHKFLSTSGDVSTCKSKASQDVNVAAYVQKQPSPL  
LIVASSRATYVPRFTSGSGSDPFTFTLSVQAEIDAVYFCCOHPTPTPSGTL  
EIKALBNHHHLEGGSLALITVQACHLELEFTRRRQRMEQECGYPRQLVA  
LYLARLSKMQVDVIRNALASFGSGDLSERLREDEQARLRLTLAAESERFVQ  
TGNDKAGANADEKLQVQLQSGPLVQSPQSITCTCFSGFLITVGVHWRQSPK  
GLEMLGVWGGANTDYNTPFTSRISINKDKSKQVFKMNSLQNDTALTYCARALTY  
YDEFAWQGGTIVTVSSGGGSGGSGGSDILOIQTSPVILSVSGRVSFSCA  
SOSIGTNIHWQORTNGSPRLILKYASESIISGIPSRSGSGSDFTLSINVSSEDI  
ADVYCCQNNMWPPTTFGACTLILKALEDPVPSGAPSDGALLERNYGTGAEPICD  
GGVSRSTRQNTWTVRLLQARQLREERYVYVGHGTPLEAOSIVFGVARSOD  
LDATWRGFYTAGDPALAYGAQDEPARIRINGALRVYVPRSSLPVGRYSLTLA  
APEAAGVERLIGHPLRLDAITGPEEBEGRLETLIGMLRATVIVPSAIPITDPN  
VGGLDPSSSIDPKKQALISALPDYASQPGKPRREDLK"

ORIGIN

Alignment Scores:  
Pred. No.: 4,51e-92 Length: 2793  
Score: 1308.00 Matches: 275  
Percent Similarity: 71.19% Conservative: 24  
Best Local Similarity: 65.48% Mismatches: 59  
Query Match: 53.58% Indels: 64  
DB: 6 Gaps: 5

US-09-596-774-6 (1-461) x A57343 (1-2793)

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysPheProGlyGluThrValIle 38  
DB 115 CAGGTAAACATGACGACGACTGCACTGAACTGAAGAAGCTGGAAGACGACCAAGATC 174  
QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58  
DB 175 TCTGCAAGGCTCTGGGTATCTTTCACCAACTATGAAATGATGAGTGGGGAAGCGCT 224  
QY 59 ProGlyGlnIleLysIleTyrPheMetGlyTyrPheAsnThrSerThrGlyLysSerThrPhe 78  
DB 235 CAGGAGAGGGGTTTAAAGTGATGGCTGATTAACCTCCACCTGAGAGTCAACATTT 294  
QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98  
DB 295 GCTGATGACTTCAAGGAGCGTGTGACTTCTCTTGGAAACCTCTGCCAACATGCTGCTAT 354  
QY 99 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 118  
DB 355 TTGCAGATCAACAACCTCAAAAGTGAAGCATGGCTTACATATTTCTGTGCAAGATGGAG 414  
QY 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 138  
DB 415 GTTTACCAAGGCTACGTTCTTCTGAGGACCAAGGACCAAGGTCACCGTTCTCTGGC 474  
QY 139 GlyIleGlySerGlyIleGlyLysSerGlyIleGlyIleSerAspIleGlnLeuThrGln 158  
DB 475 GGTGGCGTTCGTGGCGGCTGCTCCGCGGCGGTGGCTTTCAGATCCAGCTGACCCAG 534  
QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178  
DB 535 TCTCACAAAATTCCTGTCACCTTCAGTAGACAGAGGTAGCATCACTCGCAAGGCACT 594  
QY 179 GluAspValTyrAsnAlaValAlaTyrPyrGlnGlnLysProGlyGlnIleSerProLysLeu 198  
DB 595 CAGGATGTATATATCTGTGCTGTGATTCACAGAAACAGGACAACTCTTAACCT 654  
QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
DB 655 CTGATTACTCGGCATCTCCCGTACCTGAGTCCCTTCTCGCTTCACTGGCAGTGGC 714  
QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238  
DB 715 TCTGGGCGGATTTCACTTCACTTCACTGAGCAGTGTGAGAGGCTGGAAGACTGGCAGTTAT 774

QY 239 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrIleuLeuGlnIle 258  
 DB 775 TTCTGTTCAGCAACATTTTCGTACTCATTCACGTTCCGTCGGGACAAATTGGAGATC 834  
 QY 259 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278  
 DB 835 AAAGCTTAGAGCA----- 848  
 QY 279 LysValAsnSerThrThrThrIleProValLeuArgThrProSerProValHisProThr 298  
 DB 849 -----CCATCATTCACCATCCTGAGA 869  
 QY 299 GlyThrSerGlnProGlnArgProGlnuAapCysArgProArgGlySerValIleGlyThr 318  
 DB 870 GGG---CGGACGCTGCGCGCGCTGAC-----CGGACCAAGGCTGCGACCTGCGCT 920  
 QY 319 GlyLeuAapPheLeuGlnuAapProIleuLeuGlyTyrLeuLeuAapGlyIleLeuPheIle 338  
 DB 921 GGAGACTTTCACCCGTCATCGGACCGCGCGCTGGGAAACAATGAGACAGTGGCGCTA 980  
 QY 339 TyrGlyValIle-IleThrAlaLeuTyrIleuArgAlaIlePheSerArgSerAlaGluTh 358  
 DB 981 TCCGGTGCAGCGCGCTGCTGCTGCTTACCTGCGCGCGCGCACTGCA----- 1026  
 QY 358 AlaAlaAsnLeuGlnAapProAsnGlnLeuTyrAsnGlnLeuAanLeuGlyArgArgGly 378  
 DB 1027 -----TGGAAACCAAGTGCACCA-GGT----- 1046  
 QY 378 uGluTyrAapValLeuGlnIleuGlyLysArgAlaArgAapProGluMetGlyIleGlnI 398  
 DB 1047 -----GATCCGCAAGCCCTGGCGCCCG 1072  
 QY 398 nArgArgArgAapPro-GlnGlnGlyValTyrAsnAlaLeuGlnIleuAapIleuMetAlaG 418  
 DB 1073 GCAGCGCGCGGACCTGCGCGGAGATCCGCGACAGCGCGGACAGCGCGCTGCGGCC 1132  
 QY 418 IuAlaTyrSerGlnIleGlyThrIleuGlyGluArgArgGlyLysGlyHisAap 436  
 DB 1133 TGACCCCTGGCGCGCGCGGACGAGCGCTTCTGCGGAGGCGACCGGCAAGCAGC 1188

RESULT 15  
 LOCUS A57351 2793 bp DNA linear PAT 03-MAR-1998  
 DEFINITION Sequence 21 from Patent EP0739984.  
 ACCESSION A57351  
 VERSION A57351.1 GI:3713230  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 synthetic construct  
 synthetic construct  
 other sequences; artificial sequences.  
 REFERENCE  
 1 Weis,W.D., Schmidt,M. and Groner,B.D.  
 TITLE Bivalent polypeptides containing at least two domains  
 JOURNAL Patent: EP 0739984-A 21 30-OCT-1996;  
 SAN TIMORFORSCHUNGS GMBH (DE)

FEATURES  
 source  
 1..2793  
 /organism="synthetic construct"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32630"  
 /clone="PMS238-5-5"  
 /note="MOUSE; P.ABRUGINOSA; SYNTHETIC"  
 1..2679  
 /note="unnamed protein product"  
 /codon\_start=1  
 /transl\_table=11  
 /protein\_id="CAA03448.1"  
 /db\_xref="gi:3713231"  
 /translation="MKTAIAIAVALAFATVAADYKDDDKLHHHHNKLQVLOLO  
 SGPBLKPGETVKISCKASGYPTNYGNWAKQAGGLKMMGMINTSGSTPADDF  
 KGRPFSLSTAYALQIINLKSEDMATFCARWEVHGVPVWGQCTVTIVSSGGG  
 GSGGGSGGSGDIDLTQSHKFLSTSVDRVSTCKASQDVYNAVAVYQKPGSPKL

CDS

LIYASASRYTGVPERFTSGSGDPFTTISVQVEDLAVFCCQHFRTPTFGSGTKL  
 EIKALEHHHHHLEGGSLIALTAQAQCLPLETFTRHQPAGMELEQCPVQVLVA  
 LYLAARSMNOVDVIRNALTPSGSGGLPELTPRPOCALALTLAAESERFROG  
 TGNDEAGANMDELQVLOOSGPELKRGTIVKISCAAGYPTNYGNWAKQAGG  
 GLKMMGMINTSGSTPADDFKGRPFSLSTAYALQIINLKSEDMATFCARWEV  
 YHGVPVWGQCTVTIVSSGGSGGSGGSGSDIQLTQSHKFLSTSVDRVSTCKA  
 SODPVYNAVAVYQKPGQSPKLLIYASASRYTGVPERFTSGSGDPFTTISVQEDL  
 AVFCCQHFRTPTFGSGTKLEIKALEEPLRPVPSAGPADSDALLEENYPTGAEFLGD  
 GGVDFSTRGNMTVERLLOAHOLERGVYFGYHQTFLBAQASIVFGFRRASOD  
 LDALMRGFTVIGADPALAYGVQDOEPARGRIRNGALLRVVPPSSILPFGYRTSLTLA  
 APEAAGEVERLIGHPULPRUDATITPEBEGRLTILQWPLAERTVIPSALPTDPRI  
 VGGDLPDSITPKEQAISALPDVASQPEKPREEDLK"

## ORIGIN

## Alignment Scores:

Pred. No.: 4,51e-92 Length: 2793  
 Score: 1308.00 Matches: 275  
 Percent Similarity: 71.19% Conservative: 24  
 Best Local Similarity: 65.48% Mismatches: 59  
 Query Match: 53.58% Indels: 64  
 DB: 6 Gaps: 5

US-09-596-774-6 (1-461) x A57351 (1-2793)

QY 19 GlnIleGlnLeuValGlnSerGlyProGlnuLeuLysLysProGlyGluThrValIle 38  
 DB 115 CAGGTACACCTGCAGCACTCTGAGACTGAACTGAAGAGCCTGGAGAGACAGTCAAGTTC 174  
 QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58  
 DB 175 TCTGTCAAGGCGCTCGGGTATCTTTTCAAACTATGAAAGAACTGGGTGAAGAGCGCT 234  
 QY 59 ProGlnGlnIleuLysTyrMetGlyTyrIleAsnThrSerThrGlyIleSerThrPhe 78  
 DB 235 CCAGGACGSGGTATTAAGTGAATGGCTGGAATTAACCTCAGTGGAGAGTCAACATT 294  
 QY 79 AlaAapAapPheLysGlyArgPheAapPheSerLeuGlnuThrSerAlaAsnThrAlaTyr 98  
 DB 295 GGTATGACTTCAAGGAGCGGTTTGACTTCTTTGGAAACCTGCGCAACGCTGCT 354  
 QY 99 LeuGlnIleAsnAsnLeuLysSerGlnuAapMetAlaThrTyrPheCysAlaArgTyrGlu 118  
 DB 355 TTGCAGATCAACCACTCAAAAGTGAAGACATGGCTACATTTTCTGTGCAAGATGGAG 414  
 QY 119 ValTyrHisGlyTyrValProTyrTyrGlnIleGlnIleThrValIleSerSerGly 138  
 DB 415 GTTACCAAGGCTACGTTCTTACTGGGGCCAAAGGACACAGGTACCGTTCTCTG 474  
 QY 139 GlyGlyIleSerGlyGlyGlyIleSerGlyGlyIleGlyIleSerAapIleGlnLeuThrGln 158  
 DB 475 GGTGGCGGTTCTGTTGGCGGTGGCTCCGGCGGTGGCTTGTGATCCAGTCACTGACCGAG 534  
 QY 159 SerHisLysPheLeuSerThrSerValGlyAapArgValSerIleThrCysAlaLysSer 178  
 DB 535 TCTCACAAAATTCCTGTCCACTTCACTGAGGAGACAGGCTCACATCACTCGCAAGCCAGT 594  
 QY 179 GlnAapValTyrAsnAlaValAlaIleTyrGlnIleLysProGlyGlnIleSerProLysLeu 198  
 DB 595 CAGATGTGTAATATGCTGTGCTGTATCAACAGAAACAGGACAAATCTCTTAACCT 654  
 QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
 DB 655 CTGATTTACTGGCATCTCCCGGTACAGTGAATGCTTCTCGCTTCACTGCGCAGTGGC 714  
 QY 219 SerGlyProAapPheThrPheThrIleSerSerValGlnAlaGluAapLeuAlaValTyr 238  
 DB 715 TCTGGGCGGATTTCACTTCACTCAGCAGTGTGAGCGCTGAAGACTGCGCAGTTTAT 774  
 QY 239 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrIleuLeuGlnIle 258  
 DB 775 TTCTGTTCAGCAACATTTTCGTACTCATTCACGTTCCGTCGGGACAAATTGGAGATC 834  
 QY 259 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278

```
Db      835 ||||| 848
      835 AAAGCTCTAGAGCA-----
Qy      279 LysValAsnSerThrThrIhrLysProValLeuArgThrProSerProValHisProThr 298
      849 -----CCATCATCACCATCAGCTAG 869
Qy      299 GlyThrSerGlnProGlnArgProGlnuAspCysArgProArgGlySerValLysGlyThr 318
      870 GGG---CGGCAAGCTCGGCGCGCTGAC-----CGGCAAGCTCGGCGCGCTGACCTGCGGCT 920
Qy      319 GlyLeuAspPheLeuGlnuAspProLysLeuCysTyrLeuLeuAspGlyTleuPheIle 338
      921 GGAGACTTTCACCCCTCATCGCCAGCGCGGCTGGGAAACAAGTGAAGAGTGCAGCTA 980
Qy      339 TyrGlyValIle-IleThrAlaLeuTyrIleuArgAlaLysPheSerArgSerAlaGluTh 358
      981 TCCGGTGCAGCGGCTGCTGCTCTTACTGCGCGGCGGCTGCA----- 1026
Qy      358 rAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgG 378
      1027 -----TGAAACCAAGTCAGCA-GGT----- 1046
Qy      378 uGluTyrAspValLeuGlnuLysLysArgAlaArgAspProGluMetGlyGlyLysGlnG 398
      1047 -----GATCGCAAGCCCTCGGCGAGCCCG 1072
Qy      398 nArgArgArgAsnPro-GlnGlnuGlyValTyrAsnAlaLeuGlnuLysAspLysMetAlaG 418
      1073 GCAGCGGCGGCGAGCTGGGCGGAGATCCGGAGACAGCGGAGCGGCGGCTGCGGCT 1132
Qy      418 IuaIaTyrSerGlnuIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp 436
      1133 TGACCCCTGCGCGCGCGAGAGCGGCTTCTGCTCGGAGGCGGCGGCAACGAC 1188
Db
```

Search completed: April 20, 2005, 14:06:06  
Job time : 6200.38 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Comphen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 20, 2005, 07:10:49 ; Search time 735.356 Seconds

(without alignments)  
3711.128 Million cell updates/sec

Title: US-09-596-774-6

Perfect score: 2441

Sequence: 1 MAMVTLFLMLMAAKVPKQI.....ISTATKDYDALHMQLAPR 461

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV-xlh  
-O=/cgn2.1/USPTO\_epool\_h/US09596774/runat\_20042005\_075441\_12501/app\_query.fasta\_1.1230  
-DB=N\_Geneseq -QFM=fastap -SUFFIX=sp120.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blonsum62 -TRANS=human40.cdi  
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTPM=PCO -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USER=US09596774\_QCGN\_1.1.653 @runat\_20042005\_075441\_12501 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGEOQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOC  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_16Dec04:.\*  
1: geneseqn1980s:.\*  
2: geneseqn1990s:.\*  
3: geneseqn2000s:.\*  
4: geneseqn2001as:.\*  
5: geneseqn2001bs:.\*  
6: geneseqn2002as:.\*  
7: geneseqn2002bs:.\*  
8: geneseqn2003as:.\*  
9: geneseqn2003bs:.\*  
10: geneseqn2003cs:.\*  
11: geneseqn2003ds:.\*  
12: geneseqn2004as:.\*  
13: geneseqn2004bs:.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2441	100.0	1479	2	AAT05783 Leader-sc
2	1675	68.6	1452	2	AAV73324 Chimeric
3	1506.5	61.7	3058	2	AAV77137 Single ch
4	1503.5	61.6	1284	2	AAV77139 Single ch
5	1314.5	53.9	2011	2	AAQ28262 Fv(FRP5) -

6	1312.5	53.8	2070	2	AAT42036	Aat42036 Plasmid p
7	1312.5	53.8	2730	2	AAT42044	Aat42044 Plasmid p
8	1308	53.6	1862	2	AAT29413	Aat29413 DGAL4-DT-
9	1308	53.6	1919	2	AAT29412	Aat29412 DGAL4-DT-
10	1308	53.6	2214	2	AAT42038	Aat42038 Plasmid p
11	1308	53.6	2793	2	AAT42039	Aat42039 Plasmid p
12	1308	53.6	2793	2	AAT42043	Aat42043 Plasmid p
13	1308	53.6	3177	2	AAT42041	Aat42041 Plasmid p
14	1305	53.5	1692	2	AAT29409	Aat29409 scFv(FRP5)
15	1303	53.5	2233	2	AAQ28257	Fv(FRP5) -
16	1302	53.3	794	2	AAT42034	Aat42034 Plasmid p
17	1302	53.3	2718	2	AAT42042	Aat42042 Plasmid p
18	1302	53.3	2793	2	AAT42040	Aat42040 Plasmid p
19	1285.5	52.7	748	2	AAT05781	Aat05781 EcRb-2-EP
20	1276.5	52.3	748	2	AAQ28256	Fv heavy
21	1249	51.2	1920	10	ADH34695	Adh34695 CE7-speci
22	1249	51.2	6834	10	ADH34699	Adh34699 CE7-speci
23	1015.5	41.6	7654	8	ABX16565	Abx16565 Retrovira
24	1015.5	41.6	7654	10	ABX13168	Abx13168 Retrovira
25	1013	41.5	726	4	AAK30723	AAK30723 Anti-FIX/
26	1011.5	41.4	1422	2	AAT90509	Aat90509 DNA encod
27	1011.5	41.4	1545	2	AAT90510	Aat90510 DNA encod
28	994	40.7	771	4	AAC86584	Aac86584 DNA encod
29	994	40.7	771	9	ABT43523	Abt43523 scFVSA fu
30	981.5	40.2	909	2	AAT36878	Aat36878 741F8 ant
31	981.5	40.2	909	2	AAT91837	Aat91837 Anti-c-er
32	981.5	40.2	909	2	AAT78879	Aat78879 Single ch
33	981.5	40.2	909	2	AAV22598	AAV22598 DNA encod
34	981.5	40.2	909	2	AAV21796	AAV21796 741F8 ant
35	981.5	40.2	909	2	AAV63397	AAV63397 Anti-c-er
36	981.5	40.2	909	10	AAD61483	Aad61483 741F8 gfv
37	978.5	40.1	987	10	ADG32348	Adg32348 DNA encod
38	972.5	39.8	777	10	ADG32297	Adg32297 Mouse scF
39	972.5	39.8	2001	2	AAQ46088	Aaq46088 Sequence
40	963.5	33.5	987	10	ADG32350	Adg32350 DNA encod
41	962.5	33.4	1956	2	AAT90512	Aat90512 DNA encod
42	962.5	33.4	2079	2	AAT90513	Aat90513 DNA encod
43	960.5	33.3	777	10	ADG32299	Adg32299 Mouse scF
44	930	38.1	1350	3	AAT97844	Aat97844 DNA of a
45	928.5	38.0	800	3	AAA08034	AAA08034 Antibody

## ALIGNMENTS

RESULT 1	
ID	AAT05783 standard; cDNA, 1479 BP.
XX	
AC	AAT05783;
XX	
DT	16-MAR-1996 (first entry)
XX	
DE	Leader-scFv (FRP5): Iyc-2 hinge: zeta cDNA.
XX	
KW	Single chain antibody; scFv; antibody engineering; antitumor;
KW	tumor antigen binding; T-cell receptor; cytotoxic T-lymphocyte;
KW	monoclonal antibody; ecdb-2; cancer; cell targeting; gene therapy;
KW	adoptive immunotherapy; ss.
XX	
OS	Synthetic.
XX	
FT	Key
FT	Location/Qualifiers
FT	40..1426
FT	/*tag= a
FT	/product= "Leader-scFv(FRP5): Iyc-2 hinge: zeta"
PN	W09530014-A1.
XX	
PD	09-NOV-1995.
XX	
PF	20-APR-1995; 95WO-EP001494.
XX	
XX	02-MAY-1994; 94EP-00810244.

XX. (CIBA ) CIBA GEIGY AG.  
 PA  
 XX  
 PI Groner B, Moritz D;  
 XX  
 DR WPI; 1995-393085/50.  
 XX P-PSDB; AAR85508.  
 PT New bifunctional proteins for use in killing tumour cells - contg. a  
 PT tumour antigen binding domain, a hinge region and a zeta chain derived  
 PT from a T-cell antigen receptor.  
 XX  
 PS Example 1; Page 29-31; 46pp; English.  
 XX  
 CC cDNA (AAT05783) codes for a bifunctional protein (AAR85505) consisting of  
 CC single chain antibody scFv(FRP)5 directed against the tumour erbB-2  
 CC antigen, a hinge region, and a functional zeta chain obd. from a T-cell  
 CC receptor. The cDNA is expressed in host cells, esp. cytotoxic T-  
 CC lymphocytes, providing them with a defined tumour cell specificity  
 CC enabling targeting to defined tumour cells and MHC-unrestricted and MHC-  
 CC independent tumour destruction in vitro or in vivo  
 XX  
 SQ Sequence 1479 BP; 369 A; 395 C; 394 G; 321 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 3,02e-183 Length: 1479  
 Score: 2441.00 Matches: 461  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-596-774-6 (1-461) x AAT05783 (1-1479)  
 QY 1 MetAlaTTPvAlTTPThLeuLeuPheLeuMetAlaAlaAlaAlaValProLysGlnIle 20  
 DB 40 ATGGCTTGGGTGGATGACCTTGTCTATCTGTATGGAGCTGCCAAGTCCCAAGCAGATC 99  
 QY 21 GlnLeuValGlnSerGlyProGlnuLeuLysPProGlyGlnThrValLysIleSerCys 40  
 DB 100 CAGTTGGTGGAGCTGCGAGCTGAGCTGAAGAACTGAGAGACAGTCAAGATCTCTGTC 159  
 QY 41 LysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTPvAlLysGlnAlaProGly 60  
 DB 160 AAGGCTCTGGGTATCTCTTTCACAAATCATGAAATGAACTGGTGAGCAAGCTCCAGCA 219  
 QY 61 GlnGlyLeuLysTPMetGlyTyrIleAsnThrSerThrGlyGlnuSerThrPheAlaAsp 80  
 DB 220 CAGGCTTTAAGTGGATGGCTGGATTAACACTTCCTGAGAGAGTCAACATTTGCTGAT 279  
 QY 81 AspPheLysGlyATgPheAspPheSerLeuGlnuThrSerAlaAsnThrAlaTyrLeuGln 100  
 DB 280 GACTTCAAGGAGCGATTTGACTTCTTTGAAACCTCTGCCAACACTGCTATTTGCTGAG 339  
 QY 101 IleAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaAGTTPGluValTyr 120  
 DB 340 ATCAACAACCTCAAAAGAGACATGCTCATATTTCTGTGCAAGATGGAGAGTTTAC 399  
 QY 121 HisGlyTyrValProTyrTTPGlyGlnGlyThrThrValThrValSerSerGlyGlyGly 140  
 DB 400 CACGGCTACGTTCTTACTTGGGGCCAAAGGACCAAGCTCAGCTTCTCTGGCGGTGGC 459  
 QY 141 GlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGlnSerHis 160  
 DB 460 GGTTCTGTGGCGGTGGCTCCGGCGGTGGCGGTCTGACATCCAGCTGACCCAGTCTCAC 519  
 QY 161 LysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSerGlnAsp 180  
 DB 520 AAATTCCTGTCACTTCAGTAGAGACAGGCTCAGCTACCTCCCAAGGCGCAGCAGAT 579  
 QY 181 ValTyrAsnAlaValAlaTTPYrGlnGlnLysProGlyGlnSerProLysLeuLeuIle 200  
 DB 580 GTGTATATGCTGTGCTGCTGATCAACAGAAACAGAGCAATCTCTTAACATTCTGATT 639

QY 201 TyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 220  
 DB 640 TACTGGCATCTCCCGGTACATGAGTCCCTTCTGCTTCACTGGCAGTGGCTTGGG 699  
 QY 221 ProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCys 240  
 DB 700 CCGGATTTCACTTTCACCATCAGACAGTGGCTGAAGACTGGAGAGTTATTTCTGT 759  
 QY 241 GlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnLysAla 260  
 DB 760 CAGCAACATTTCTGACTCCATTCACGTTGGCTGGGCAAAATTTGGAGATCAAGCT 819  
 QY 261 LeuGlnLysSerAsnSerValMetTyrPheSerSerValValProValLeuGlnLysVal 280  
 DB 820 CTAAAGATCAACAACCTGGTATGATCTTCAATGTTCTGTGCGCAAGTCTTCAGAAAGTG 879  
 QY 281 AsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThrGlyTyr 300  
 DB 880 AACTTACTACTACCAAGCCAGTCTGCAATCTCTCACTGACCTTACCGGAGACA 939  
 QY 301 SerGlnProGlnArgProGlnuAspCysArgProArgGlySerValLysGlyThrGlyLeu 320  
 DB 940 TCTCAGCCCAAGACCAAGAAATTTGCGGCCGCTGCTCAGTGAAGGGGACCGGATTG 999  
 QY 321 AspPheLeuGlnuAspProLysLeuCysTyrLeuLeuAspGlyLysLeuPheIleTyrGly 340  
 DB 1000 GACTTCTTAAGAGATVCCCAAACTCTGCTACTGTGATGTAATCTCTTCACTTACCGGA 1059  
 QY 341 ValIleIleThrAlaLeuTyrTyrLeuArgAlaLysPheSerArgSerAlaGlnuThrAla 360  
 DB 1060 GTCATATCAAGCGCTGTACTCTGAGAGCAAAATTCAGCAGAGTGCAGAGACTGCTGCC 1119  
 QY 361 AsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlnLysArgGlnuTyr 380  
 DB 1120 AACCTGAGAGACCCCAACAGCTCTCAATGAGCTCAATCTAGGCGAAGAGAGAAATAT 1179  
 QY 381 AspValLeuGlnLysLysAlaArgAlaArgAspProGlnuMetGlyGlyLysGlnuArgArg 400  
 DB 1180 GAGCTCTTGAAGAAAGAGCGGCTCGGATTCAGAGATGGAGGCAAAACGACAGAGAG 1239  
 QY 401 ArgAsnProGlnuGlnuGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlnuAlaTyr 420  
 DB 1240 AGGAACCCCAAGAGAGCGTATCAATGCACTGCAGAAAGACAATGGACAGAGCTTAC 1299  
 QY 421 SerGlnuIleGlyThrLysGlyLysArgArgArgGlyLysGlyHisAspGlyLeuTyrGln 440  
 DB 1300 AGTGAGATCGGCACAAAGGCGAGAGCGGAGAGGCAAGGCGCACATGGCTTTTACAG 1359  
 QY 441 GlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThrLeuAlaPro 460  
 DB 1360 GGTCTCAGACCTGCCACCAAGACACCTTGAATGCCCTGCATATGCAAGACCTGCCCT 1419  
 QY 461 Arg 461  
 DB 1420 CGC 1422  
 RESULT 2  
 ID AAV73324 standard; DNA; 1452 BP.  
 XX  
 AC AAV73324;  
 XX  
 DT 26-FEB-1999 (first entry)  
 XX  
 DE Chimeric CD19/CTCR DNA.  
 XX  
 KW Single chain; antibody; B-cell marker; CD19; scFvCD19; gene therapy;  
 KM retroviral vector; tumour-associated antigen; cancer; immunoglobulin;  
 KM CD8 alpha chain; hinge region; theta chain; transmembrane domain;  
 KM intracellular domain; T-cell receptor-CD3 complex; ss.  
 XX  
 OS Homo sapiens.





OY 459 ALAPROARG 461  
 DB 1441 GCCCCTCGC 1449

RESULT 3  
 AAT77137  
 ID AAT77137 standard; DNA; 3058 BP.  
 AC AAT77137;  
 XX  
 DT 25-MAR-2003 (revised)  
 DT 04-MAR-1998 (first entry)

Single chain antigen hybrid receptor DNA.

Hybrid receptor; single chain antigen; gene therapy; diagnosis;  
 signal conduction; receptor; control region; ss.

Synthetic.  
 Homo sapiens.

Key  
 misc\_signal  
 Location/Qualifiers  
 /tag= a  
 /note= "Noted in specification"  
 /tag= b  
 /note= "This sequence contains an intron"  
 /tag= d  
 /label= leader\_sequence  
 /note= "This sequence contains an intron"  
 /tag= c  
 /tag= e  
 /product= "single\_chain\_antigen\_hybrid\_receptor"  
 /tag= f

mat\_peptide  
 862..2133  
 /tag= c

polya\_site  
 2514..2519  
 /tag= f

W09720938-A2.  
 12-JUN-1997.  
 03-DEC-1996; 96WO-DE002334.  
 05-DEC-1995; 95DE-01045351.  
 (UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.  
 Mertelsmann R, Kulmburg P, Rosenthal F;  
 WPI; 1997-319784/29.  
 P-PSDB; AAW24025.

Cells with hybrid receptor having extracellular and intracellular regions  
 of different origins - useful in gene therapy and diagnosis of tumours.

Example 3; Fig 4; 46pp; German.

This DNA sequence encodes a novel single chain antigen hybrid receptor  
 (HR) and contains an extracellular domain specific for the hapten 4-  
 hydroxy-5-iodo-3-nitrophenyl acetate (NIP) coupled to the CD8-alpha  
 region and the transmembrane and signal-conducting intracellular parts of  
 the CD3-zeta molecule. Such hybrid receptors comprise a receptor part  
 localised on the outside of the cell and specific to a particular signal  
 molecule and a receptor part originating from another receptor, localised  
 on the inside of the cell and capable of setting off a signal inside the  
 cell. The cell should also contain at least one other gene construct with  
 a control region which can interact with the signal sent out by the  
 hybrid receptor and thereby control expression of a transgene bound to  
 this control region. Such cells are useful in gene therapy or for

CC diagnostic purposes. (Updated on 25-MAR-2003 to correct PR field.)  
 XX Sequence 3058 BP; 689 A; 831 C; 853 G; 685 T; 0 U; 0 Other;

Alignment Scores:  
 Pred. No.: 4,43e-109 Length: 3058  
 Score: 1505.50 Matches: 299  
 Percent Similarity: 75.27% Conservative: 45  
 Best Local Similarity: 65.43% Mismatches: 86  
 Query Match: 61.72% Indels: 27  
 DB: 2 Gaps: 6

US-09-596-774-6 (1-461) x AAT77137 (1-3058)

OY 9 PheLeuMetAlaAlaAlaValProLysGlnIleGlnLeuValGlnSerGlyProGlu 28  
 DB 832 TTTCCTCTTCTCTCCACAGGTGTCACCTCCAGGTCACCTGACGACGTCGTGGGCTGAG 891

OY 29 LeuLysLysProGlyGlnThrValLysIleSerCysLysAlaSerGlyTyrProPheThr 48  
 DB 892 CTTGTGAAGCTGGGGCTTCAGTGAACCTGCTCCGCAAGCTTCTGCTACACCTTACC 951

OY 49 AsnTyrGlyMetAsnTyrValLysGlnAlaProGlyGlnIleuLysTyrMetGlyTyr 68  
 DB 952 AGCTACTGAGATGCACTGGGTGAAGCAAGAGCGCTGACGAGCTTGATGGAAG 1011

OY 69 IleAsnThrSerThrGlyGlnSerThrPheAlaAspPheLysGlyArgPheAspPhe 88  
 DB 1012 ATTGATCTTAATGATGGGTGGTACTCAAGTACATGAGAGTCAAGAGCAAGCCACACTG 1071

OY 89 SerLeuGlnThrSerAlaAsnThrAlaTyrLeuGlnIleAsnAsnLysSerGlyAsp 108  
 DB 1072 ACTGTAGACAAACCTCCACACAGCGCTTCATGACAGTCAAGCCTGATCTGAGAC 1131

OY 109 MetAlaThrTyrPheCysAlaArgTyrGluValTyrHisGlyTyrValProTyrTyr 127  
 DB 1132 TCTGGCTCTAATGATGTCAGATGCAATGATGATGATGATGATGATGATGATGATG 1191

OY 128 GlyGlnGlyThrThrValThrValSerSerGlyGlyGlySerGlyGlyGlySer 147  
 DB 1192 GGCCAGGAGCCAGCGTACCGTCTCTCTGAGTGGAGGCGGTGACGAGGAGTGGCTCT 1251

OY 148 GlyGlyGlyGlySerAspIleGlnLeuThrGlnSerHisLysPheLeuSerVal 167  
 DB 1252 GCGGCTGGCGGATCTCAGGCTGTGGACACAG---GAATCTGACATCACCATCACT 1308

OY 168 GlyAspArgValSerIleThrCysLysAlaSerGlnAspValTyr-----AsnAla 184  
 DB 1309 GTTGAACAGTCCACTCACTCTGCTGCTCAAGTACGAGGCTGTTCACATGATAC 1368

OY 185 ValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSerAlaSer 204  
 DB 1369 GCCAAGTGGGTCCAAAGAAACCAAGATCTTATTCATGCTGCTAATAGTGTACCAAC 1428

OY 205 SerArgTyrThrGlyAlaProSerArgPheThrGlySerGlySerGlyProAspPheThr 224  
 DB 1429 AACCGAGCTCCAGGTGCTCTGCTGCAATTCACAGCTCCCTGATGCAACAAGCTGCC 1488

OY 225 PheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGlnHisPhe 244  
 DB 1489 CTCACATCACAGGGGACAGTACAGTGAAGGCAATATTTCTGCTCTTATGATGAC 1548

OY 245 ArgThrProPheThrPheGlySerGlyThrLysLeuGlnIleLysAlaLeuGlnIleSer 264  
 DB 1549 AGCAACATTTGGGTGTTCGTGGAGGAACAACATGACATGTC-----CTCAG----- 1596

OY 265 AsnSerValMetCysPheSerSerValValProValLeuGlnLysValAsnSerThrThr 284  
 DB 1597 -----GAATCTACTACT 1608

OY 285 ThrLysProValLeuArgThrProSerProValHisProThrGlyThrSerGlnProGln 304  
 DB 1609 ACCAGCCAGTGTGGAACTCTCACTGACACCTTACCGGAGACATCTCAGCCCCAG 1668

Qy	305	ArGPvogiuaapCyBaRgProaRgiYseRVallylsGlyYthRnglyVleuAspPheUenU	324
Db	1669	AGACCAAGAAATTGTCGGCCCGCTGCAGTGAAGGGGACCGGATTGGACTTC-----	1722
Qy	325	AspProLyvLeuCyBryrVleuLeuAspGlylleuPheileYrGlyValillethr	344
Db	1723	GACCCCAAACTCTGCTACTTCTAGATGGATCCCTTCCTACCTTCAGAGTCAATCA	1782
Qy	345	AlAleuTyrlleuAargAlalySpHeserArgseralagIuThrAlaIAsnUenUAsp	364
Db	1783	GCCCTGTCTCCGAGACGAAAAATTCAGCAGAGTGCAGAGCTGGCCAACTCGAGAC	1842
Qy	365	ProAenGnlleuTyrAsnGnlleuAsnUenUgIyArGAgGlyUgnIuTyrAspValleuU	384
Db	1843	CCCAACCGCTCTACATGAGTCAATAGGCGCAGAAAGAGAAATAGAGCTTGGAG	1902
Qy	385	LyblybAargAlaArGAspProGlnUmetGlyglYvpsGIngnIaRgArGAspProGln	404
Db	1903	AAGAAGCGGGCTCGGAGTCCAGAGATGGAGGCAAAACAGCAGAGAGGAGAACCCCA	1962
Qy	405	GlUgnIyValTyrrAsnAlaLeuGlnUlyAspLySmetAlaGluAlaTySserGlnUlleY	424
Db	1963	GAAAGCATATVCAATGCACTCAGAAACAACAATGAGAAAGCCTTACAGTGAATCGGC	2022
Qy	425	ThrylsGlyglUaRgArGArgGlyVlypsGlyHIsaRpsGlyLeuTyrgInGlyleuSerThr	444
Db	2023	ACAAAGGCGAGAGGCGGAGAGGCAAGGGGCCAGTGGCTTTACAGAGCTTCGACACT	2082
Qy	445	AlaThrlybAspThrTyrrAspAlaLeuHIsmetGlnThrleuAlaProArg	461
Db	2083	GGCAACCAAGACACCTATGATGCCCCCTGGATATGCAGACCTCGGCCCTTCGC	2133

XX	AC	AT77139	standard; DNA; 1284 BP.
XX	AT77139;		
XX	25-MAR-2003	(revised)	
XX	04-MAR-1998	(first entry)	
XX	Single chain antigen hybrid receptor partial DNA sequence.		
XX			
KW	Hybrid receptor; single chain antigen; gene therapy; diagnosis;		
XX	signal conduction; receptor; control region; ss.		
XX			
OS	Synthetic.		
OS	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	CDS	1..1276	
FT		/*tag= a	
FT		/note= "partial hybrid receptor"	
FT	sig_peptide	1..12	
FT		/*tag= b	
FT		/note= "partial leader sequence"	
FT	mat_peptide	13..1285	
FT		/*tag= c	
FT		/product= "hybrid_receptor"	
FT		/note= "single chain antigen hybrid receptor"	
XX			
PN	W09720938-A2.		
XX			
PD	12-JUN-1997.		
XX			
PF	03-DEC-1996;	96WO-DE002334.	
XX			
PR	05-DEC-1995;	95DE-01045351.	
XX			
PA	(UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.		
XX			
XX	Mertelsmann R, Kulmburg P, Rosenthal F;		

XX WPt\_1997-319784/29.  
DR P-PsDB; AAMZ4027.

PX  
PR Cells with hybrid receptor having extracellular and intracellular regions  
of different origins - useful in gene therapy and diagnosis of tumours.

PS Example 3; Fig 4; 46dp; German.

CC This partial DNA sequence encodes a novel single chain antigen hybrid  
receptor (HR). The full length receptor sequence can be found in AAU77137

CC and contains an extracellular domain specific for the hapten 4-hydroxy-5-

CC iodod-3-nitrophenyl acetate (NIP) coupled to the CD8-alpha region and the

CC transmembrane and signal-conducting intracellular parts of the CD3-zeta

CC molecule. Such hybrid receptors comprise a receptor part localised on the

CC outside of the cell and specific to a particular signal molecule and a

CC receptor part originating from another receptor, localised on the inside

CC of the cell and capable of setting off a signal inside the cell. The cell

CC should also contain at least one other gene construct with a control

CC region which can interact with the signal sent out by the hybrid receptor

CC and thereby control expression of a transgene bound to this control

CC region. Such cells are useful in gene therapy or for diagnostic purposes.  
(Updated on 25-MAR-2003 to correct PR field.)

SO Sequence 1284 BP; 334 A; 342 C; 355 G; 253 T; 0 U; 0 Other;

XX XX

[illegible]



```

QY 55 VallysglnAlaProGlyGlnGlyLeuLeuYrTrpMetGlyTrpIleAsnThrSerThrGly 74
DB 205 GTGAAGCAGGCTCCAGACAGAGGTTTAAAGTGAAGGCTGATTAACCTCCCTCAGCA 264
QY 75 GluSerThrPheAlaAspAspPheLeuGlyArgPheAspPheSerLeuGlnThrSerAla 94
DB 265 GAGTCAACATTGTCGATGACTTCAAGGAGCGTTTGACTTCTTTGGAAACCTCTGCG 324
QY 95 AsnThrAlaTrpLeuGlnIleAsnAsnLeuLeuSerGluAspMetAlaThrTrpPheCys 114
DB 325 AACACTGCTATTTCAGATGAAACCACTCAAAAGTGAAGCATGCTACATATTCTGCT 384
QY 115 AlaArgTrpGluValTrpPheGlyTrpValProTrpGlyGlnGlyThrThrValThr 134
DB 385 GCAGATGGAGGTTTACACAGGCTACGTTCTTCTGAGGCGCAAGAGACAGGCTACAC 444
QY 135 ValSerSerGlyGlyGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlySerAlaPhe 154
DB 445 GTTTCCTCTGGCGGTGGCGGTCTGTGGCGGTGGCGGTGGCGGTGGCGGTCTGACATC 504
QY 155 GlnLeuThrGlnSerHisLeuPheLeuSerThrSerValGlyAspArgValSerIleThr 174
DB 505 CAGCTGACCCAGTCTCACAAATTCCTGCTCAGTACAGAGACAGGCTCAGCATCAC 564
QY 175 CysLeuValSerGluAspValTrpAsnAlaValAlaTrpTrpGlnGlnLeuAspProGlyGln 194
DB 565 TGCAAGGCGCAGTACAGATGTATATGCTGTGCTGATCAACAGAAACAGAGACAA 624
QY 195 SerProLeuLeuLeuIleTrpSerAlaSerSerArgTrpThrGlyValProSerArgPhe 214
DB 625 TCTCTTAACCTTGATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTA 684
QY 215 ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 234
DB 685 ACTGGCAGTGGCTCGGCGCGGCTTCACTTCAACATCAGCAGTGGCAGGCTCAAGAC 744
QY 235 LeuAlaValTrpPheCysGlnGlnHisPheArgTrpPheThrPheGlySerGlyThr 254
DB 745 CTGGAGATTATTCTGTGACAGAACATTTCTGCTTCACTTCACTTCACTTCACTTCACT 804
QY 255 LysLeuGlnIleLeuAlaLeuGlnIleSerAsnSerValMetTrpPheSerSerVal 274
DB 805 AAATTGGAGATCAAGGCTTCAAG----- 828
QY 275 ProValLeuGlnLysValAsnSerThrThrTrpLysProValLeuArgThrProSerPro 294
DB 829 -----GGCGGACAGCTGCGCGGCTGAC-CGCGCAGCAGGCGCTG 866
QY 295 ValHisProThrGlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySer 314
DB 867 CCACCTCCCGCTGAGACTTTCACCCGTCATCGCA-----GCCGCG 908
QY 315 ValLysGlyThrGlyLeuAspPheLeuGluAspProLysLeuGlyTrpLeuLeuAspGly 334
DB 909 CGGCTGGGAACA-----ACTGGA 926
QY 335 IleLeuPheIleTrpGlyValIle-IleThrAlaLeuTrpLeuArgAlaLysPheSer-- 353
DB 927 GCAGTGGCGCTATCCGTCGACAGCGGCTGCTCCCTCTACCTGCGGCGCAGCTGTCATG 986
QY 354 -----ArgSerAlaIleThrAla----- 359
DB 987 GAACCAAGTCGACGAGTGATCCGAAACCCCTGGCGGAGCCCGCGCAGCGCGGACCT 1046
QY 360 -----AlaAsnLeuGlnAspProAsnGlnLeuTrpAsnGlnLeuAsnLeuGlyArgArg 377
DB 1047 GGGCGAAGCGATCCCGCAGACCGCGAGCAGCCCGCTGCGCTGACCTT-GGCGCGCG 1105
QY 377 gGlnGluTrpAspValLeuGlnLysValArgAlaArgAspProGluMetGlyGlyGly 397
DB 1106 CCGAGAG-----CGAGCGCTGCTGCTCG-----GCAGGCGCAGCG 1138

```

```

QY 397 nglnArgArgAsnProGlnGlnGlyValTrpAsnAlaLeuGlnLysAspLysMetAl 417
DB 1139 GCAAGCAGCA----- 1148
QY 417 aglnAlaTrpSerGluIleGlyThrLysGlyGlnArgArgArgGlyLys 433
DB 1149 -----GGCGGCGCGCGCAACGCGCAGCTGTGTAG 1178

RESULT 6
AAT42036
ID AAT42036 standard; DNA; 2070 BP.
AC AAT42036;
XX 29-JAN-1997 (first entry)
DT
DE Plasmid pSW202-5 fragment encoding scFv(FRP5)-ETA.
XX
KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pSW202-5; cancer; therapy;
XX antitumour; exotoxin A; ETA; ss.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..1956
FT sig_peptide 1..63
FT /tag= a
FT /note= "Omiga signal peptide"
FT mat_peptide 64..1953
FT /tag= c
FT /product= "scFv(FRP5)-ETA"
XX
PN EP739984-A1.
XX
PD 30-OCT-1996.
XX
PF 26-APR-1995; 95EP-00106275.
XX
PR 26-APR-1995; 95EP-00106275.
XX
PA (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
PI Weis W, Schmidt M, Groner B;
XX
DR MPI; 1996-478748/48.
XX
DR P-PSDB; AAM05156.
XX
PT Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX
PS Example 7; Page 21-22; 52pp; English.
XX
CC A fragment (AAT42036) of plasmid pSW202-5 codes for scFv(FRP5)-ETA
CC (AAM05136), comprising the single-chain binding region of murine
CC monoclonal antibody FRP5, which is specific for human epidermal growth
CC factor receptor erbB-2, joined to exotoxin A (ETA). A fragment of plasmid
CC pMW52-5 (see also AAT42036) encoding scFv(FRP5) (AAM05134) was ligated
CC to pSW202-5, a plasmid contg. the Pseudomonas aeruginosa PAK ETA gene, to
CC produce pSW202-5. This can be utilised in the construction of bacterial
CC expression vectors (see also AAT42037-44) encoding bivalent fusion
CC proteins (AAM05137-44) useful as antitumour agents
XX
SQ Sequence 2070 BP; 402 A; 675 C; 628 G; 365 T; 0 U; 0 Other;

Alignment Scores: 5.74e-94 Length: 2070
Pred. No.: 1312.50 Matches: 284

```

Percent Similarity: 68.59% Conservative: 13  
 Best Local Similarity: 65.59% Mismatches: 55  
 Query Match: 53.77% Indels: 83  
 DB: 2 Gaps: 8

US-09-596-774-6 (1-461) x AAT42036 (1-2070)

```

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 38
DB 115 CAGGTACCACTCCAGACAGCTGAGCCTGAACTGAAAGACCTGGAGAGACAGTCAAGATC 174
QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 58
DB 175 TCCTGCAGAGCCTCTGGGTATCCTTCAACAATATGAAATGAACTGGGTGAAGCAGGCT 234
QY 59 ProGlyGlnGlyLeuLysTyrMetGlyTyrPheLeuThrSerThrGlyLysThrPhe 78
DB 235 CCAGGACAGGCTTTAAAGTGAAGGCTGGATTAAACCTCCACTGGAGAGTCAACATTT 294
QY 79 AlaAspAspPheLysGlyLysArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
DB 295 GCTGATGACTTCAAGGAGCGTTTGACTTCTCTTGGAAACCTCTGCCAACACTGCCAT 354
QY 99 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
DB 355 TTGCAGATCAACACCTCAAAAGTGAAGACATGCTACATATTTCTGCAAGATGGAG 414
QY 119 ValTyrHisGlyTyrValProTyrTyrProGlyGlnGlyThrThrValThrValSerSerGly 138
DB 415 GTTTACCAAGCGCTACGCTTCTTACTGGGGCCAGGAGACACGCTACCGTTCCTGGC 474
QY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158
DB 475 GGTGGCGGCTTCTGGGGCGGTGCTCCGCGGTGGCTTGAATCATCACTGACCTGAC 534
QY 159 SerHisLysPheLeuSerThrSerValGlyAspGlyValSerIleThrCysLysAlaSer 178
DB 535 TCTCACAATTCCTGCTCACTTCACTAGAGAGACAGGTCACATCCTCGCAAGCCAGT 594
QY 179 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 198
DB 595 CAGAGTGTATATGCTGCTGCTGCTGATCAACGAAACAGACATCTCTTAACCT 654
QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
DB 655 CTGATTACTCGGCATCTCCCGGACACTGAGTCCCTTCCTCGCTCAGCGCAGTGC 714
QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
DB 715 TCTGGGCGGATTCATCTTCACTCACTGAGAGTGCAGGCTGAAAGCTGGCAGTTAT 774
QY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrLysLeuGluIle 258
DB 775 TTCGTGACGAAACATTTTCGATCCATTCACGTTCCGCTGGGAGCAAAATTTGGAATC 834
QY 259 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 278
DB 835 AAAGCTTAGAGCA----- 848
QY 279 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 298
DB 849 -----CCATATCAACATCACTAGA 869
QY 299 GlyThrSerGlnProGlnArgProGluAspCysArgProArgIlySerValLysGlyThr 318
DB 870 GGG-----CGGACGCTGGCGCGCGTAC-----CGCGACACAGCCCTGCCACCTGCCGCT 920
QY 319 GlyLeuAspPheLeuGluAspProLysLeuGlyTyrLysLeuAspGlyIleLeuPheIle 338
DB 921 GGAGACTTTTCAACCGTCATCGCCAGCGCGCGGAGAAACATGAGAGAGATGCGGCTA 980
QY 339 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSer----- 353
  
```

```

DB 981 TCCGATGAGCGGCTGTCGCCCTTCACTGAGCGCGGCACTGTATGAAACAGTCTGA 1040
QY 354 -----ArgSerAlaGluThrAla-----AlaAs 361
DB 1041 CCAGGTGATCCGCAACGCTCGGACCGCCGAGCGGCGGACCTGGGCGGAAGCGAT 1100
QY 361 nLeuGlnAspProAsnGlnLeuTyrAsnAlaLeuAsnLeuGlyLysArgGlyGluTyrAs 381
DB 1101 CCGGAGAGCGCGGAGCAGCGCCGCTGCGCTGACCTT-GGCGCGCGCGAGG----- 1154
QY 381 pValLeuGluLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGlnArgArg 401
DB 1155 -----CGAGGCTTGTCTCG-----GCAAGGACCGGCAACAGACA-- 1190
QY 401 GAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSe 421
DB 1190 ----- 1190
QY 421 rGluIleGlyThrLysGlyGluArgArgArgLys 433
DB 1191 -----GGCGGCGCGGCAACGCGAGCGTGTAG 1220

RESULT 7
AAT42044
ID AAT42044 standard; DNA; 2730 BP.
XX
AC AAT42044;
XX
XX 29-JAN-1997 (first entry)
DT
DE Plasmid pMS246-5-5 encoding scFv2(FRP5)-ETA (version 3).
XX
KW Single chain antibody; scFv; monoclonal antibody; Mb; EGF; etbB-2;
KW epidermal growth factor; receptor; plasmid pMS246-5-5; cancer;
KW exotoxin A; ETA; antitumour; ss.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..2700
FT sig_peptide 1..63
FT mat_peptide /note= "OmpA signal peptide"
FT /tag= c
FT /product= "scFv2 (FRP5/FRP5)-ETA"

EP739984-A1.
PD 30-OCT-1996.
XX
PF 26-APR-1995; 95EP-00106275.
XX
PR 26-APR-1995; 95EP-00106275.
XX
PA (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX
DB WPI: 1996-478748/48.
DB P-PSDB; AAM05144.
XX
PT Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX
PS Example 12; Page 43-45; 52pp; English.
CC A fragment (AAT42044) of bacterial expression plasmid pMS246-5-5 codes
  
```

CC for scfv2(FRPS/225)-ETA (AAW05144), and is obtd. by ligating 2 copies of  
 CC DNA coding for the single-chain binding region of murine monoclonal  
 CC antibody FRP5 (see also AAT2034) and portions of *Pseudomonas aeruginosa*  
 CC exotoxin A. pMS246-5-5 can be utilised in the produ. of bivalent fusion  
 CC protein in bacterial (esp. *E. coli*) host cells. Such fusion proteins (see  
 CC also AAW05138-43) bind the epidermal growth factor receptor and are  
 CC useful as antitumour agents

XX  
 SQ Sequence 2730 BP; 567 A; 832 C; 798 G; 533 T; 0 U; 0 Other;

# Alignment Scores:

Pred. No.:	8.02e-94	Length:	2730
Score:	1312.50	Matches:	284
Percent Similarity:	68.59%	Conservative:	13
Best Local Similarity:	65.59%	Mismatches:	55
Query Match:	53.77%	Indels:	83
DB:	2	Gaps:	8

US-09-596-774-6 (1-461) x AAT2044 (1-2730)

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuValysProGluValThrValylsle 38  
 Db 115 CAGGTACCACTGACACAGCTGACCTGAACTGAAGAACCTGGAGAGCACTCAAGATC 174  
 QY 39 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValylsGlnAla 58  
 Db 175 TCCTCCAAAGCCTCTGGGTATCTTCCAAACTATGAACTAGCGGTGAAGCAGGCT 234  
 QY 59 ProGlyGlnGlyLeuValTyrPheMetGlyTyrIleAsnThrSerThrGlyIleSerThrPhe 78  
 Db 235 CCAGACACGGGCTTAAAGTGAAGTGGCTGATTAACCTCCACCTGAGAGTCAACATTT 294  
 QY 79 AlaAspAspPheLeuValYargPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98  
 Db 295 GCTGATGACTTCAAGAGACGGTTTACCTCTTGGAAACCTCTGGCAACACCTGCTAT 354  
 QY 99 LeuGlnIleAsnAsnLeuValysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 118  
 Db 355 TTGCAGATCAACAACCTCAAAAGTGAAGCATGGCTACATATTCTGTGCAAGATGGAG 414  
 QY 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyValThrValThrValSerGly 138  
 Db 415 GTTACACAGCGCTACGTTCTTACTGAGGCGCAAGGAGACCACTGCTCTCTGCGC 474  
 QY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerGlyGlyGln 158  
 Db 475 GGTGGCGGTCTGGTGGCGGTGCTCCGGCGGTGGCGGTCTGACATCCAGCTGACCGAG 534  
 QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer 178  
 Db 535 TCTCAAAATTCCTCTCCACTTCAAGAGACAGGCTCAGATCACTGCAAGGCCACT 594  
 QY 179 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 198  
 Db 595 CAGGATGTGTAATATGCTGTTGCTGATCAACAGAAACAGCAATCTCTAAACTT 654  
 QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
 Db 655 CTGATTACTCGGCATCTCTCCCGGTACACTGAGAGTCCCTTCTGCTTCACTGGCAGTGGC 714  
 QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238  
 Db 715 TCTGGGCGGATTTCACTTTCACATCAGCACTGAGAGCTGAAAGACTGGCAGTTTAT 774  
 QY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrHisLeuGluIle 258  
 Db 775 TTCTGTCAAGCAATTTTGTGACTCATTCACGTTCCGCTGGGAGCAAAATTTGAGATC 834  
 QY 259 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278  
 Db 835 AAAGCTTAGAGCA----- 848  
 QY 279 LysValAsnSerThrThrLysProValLeuArgThrProSerProValHisProThr 298

Db 849 -----||||| 869  
 QY 229 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValylsGlyThr 318  
 Db 870 GGG---CGGCGAGCTGGCGCGCTGAC-----CGGAGCAAGGCGCTGCCACTTCCGCT 920  
 QY 319 GlyLeuAspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyIleuPheIle 338  
 Db 921 GGAGACTTTCACCCCTCATCTGCCAGCCCGCGCGCTGGGAAACACTGAGACACTGGCGCTA 980  
 QY 339 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSer----- 353  
 Db 981 TCCGGTGCAGGCGGTGTCGCGCCCTTACTGCGCGCGGCACTGATGAAACAGATCGA 1040  
 QY 354 -----ArgSerAlaGluThrAla-----AlaAs 361  
 Db 1041 CCAGGTATCCGCAACGCCCTGGCCAGCCCGCGAGCGGCGGCACTGGGCGAAGCGAT 1100  
 QY 361 nLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGluYargArgGluTyrAs 381  
 Db 1101 CCGGAGACGCCGAGAGCGGCCCTGCTGCGCTTACCTT-GGCCCGCGGAGAG----- 1154  
 QY 381 pValLeuGluLysLysArgAlaArgAspProGluMetGlyLysGlnGlnArgArg 401  
 Db 1155 -----CGAGGCGCTTCGTCCG-----GAGGGGACCGGCAACGAGCA-- 1190  
 QY 401 GAsnProGlnGluValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSe 421  
 Db 1190 ----- 1190  
 QY 421 rGluIleGlyThrLysGlyGluArgArgGlyLys 433  
 Db 1191 -----GGCGGCGCGGCAACGCCGAGCTGCTGAG 1220

RESULT 8  
 AAT29413  
 ID AAT29413 standard; DNA; 1862 BP.  
 XX  
 AC AAT29413;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 19-AUG-1996 (first entry)  
 XX  
 DE DGAL4-DT-scfv(FRPS) gene in pSM5-GDS.  
 XX  
 KW Nucleic acid transfer system; gene transfer; gene therapy;  
 KW cell targeting; multidomain protein; vector; cancer; GAL4;  
 KW diptheria toxin; single chain antibody; scfv; FRP5; ss.  
 XX  
 OS Saccharomycetes; cerevisiae.  
 OS Corynebacterium; diptheriae.  
 OS Mus sp.  
 OS Chimeric.  
 OS Synthetic.

Key Location/Qualifiers  
 CDS 1..1908  
 FT /\*tag=a  
 FT /product="GAL4-DT-scfv(FRPS) fusion protein"  
 XX  
 PD MO9613599-AI.  
 PD 09-MAY-1996.  
 XX  
 PF 31-OCT-1995; 95WO-EP004270.  
 XX  
 PR 01-NOV-1994; 94EP-00810627.  
 XX  
 PA (WELLS/) WELLS W.  
 XX  
 PI Weis W, Fominaya J;

DR WPI; 1996-239505/24.  
 DR P-PSDB; AAR95057.  
 PT Nucleic acid transfer system for gene therapy, e.g. against cancer -  
 PT includes toxin translocation domain to target nucleic acid to specific  
 cell.  
 XX  
 XX Example 20; Page 84-86; 106pp; English.  
 CC Plasmid pSW5-GD5 contains a fusion gene (AAT29413) coding for a  
 CC multidomain fusion protein (AAR95057) a FLAG epitope, amino acids 2-147  
 CC of yeast GAL4, amino acids 196-384 of diphtheria toxin (DT) and a single  
 CC chain antibody, scFv(FRP5), based on a monoclonal antibody raised against  
 CC the HER2 antigen of human tumour cells, all separated by spacer peptides.  
 CC The GAL4 moiety acts as a DNA binding domain, DT as a translocation  
 CC domain, and scFv(FRP5) as a ligand domain. The multidomain protein, which  
 CC can be expressed in E. coli, is used with an effector nucleic acid in a  
 CC novel nucleic acid transfer system suitable for gene therapy. The  
 CC effector nucleic acid includes the gene to be delivered and a cognate  
 CC structure for the GAL4 DNA binding domain. (Updated on 16-OCT-2003 to  
 CC standardise OS field)  
 CC  
 SQ Sequence 1862 BP; 522 A; 444 C; 446 G; 450 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,14e-93 Length: 1862  
 Score: 1308.00 Matches: 245  
 Percent Similarity: 97.64% Conservative: 3  
 Best Local Similarity: 96.46% Mismatches: 6  
 Query Match: 53.58% Indels: 0  
 DB: Gaps: 0  
 US-09-596-774-6 (1-461) x AAT29413 (1-1862)  
 QY 19 GlnIleGlnLeuValGlnSerGlyProGlyLeuLysLysProGlyGluThrValLysIle 38  
 DB 1090 CAGGTACACAGCTGACAGCTGAGCTGACCTGAAAGAGCCGAGAGACAGTCAAGATC 1149  
 QY 39 SerCysValSalaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 58  
 DB 1150 TCCTGCAAGGCGCTCGGGTATCTTTCACAACTATGAAATGAACTGGGTGAACAGAGCT 1209  
 QY 59 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 78  
 DB 1210 CCAGGACAGGGTTTAAAGTGAAGTGGCTGATTAACACCTCAGCTGAGAGTCAACATTT 1269  
 QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98  
 DB 1270 GCTGATGACTTCAAGGAGCGGTTTGACTTCTTTGGAAACCTCTGCCACAGCTGCTAT 1329  
 QY 99 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 118  
 DB 1330 TTGCAGATCAACCACTCAAAAGTGAAGACATGCTCATATTTCTGCAAGATGGAG 1389  
 QY 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyIleThrValThrValSerSerGly 138  
 DB 1390 GTTACCAAGCGCTACGTTCTTACTGGGGCCAAAGGACACAGGTCACCGTTCTCTGGC 1449  
 QY 139 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 158  
 DB 1450 GGTGGCGGTTCTGGTGGCGGTGCTCCGGCGGTGCTGACATCCAGTCAACCCAG 1509  
 QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysValSalaSer 178  
 DB 1510 TCTCAAAATTCCTGCTCCTCCTCAGTAGAGACAGGGTCAACATCACCCTCAAGGCCAGT 1569  
 QY 179 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 198  
 DB 1570 CAGGATGTGATATGCTGTTGCTGTGATCAAGAAACAGACAAATCTCTAAACTT 1629  
 QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
 DB 1630 CTGATTACTCTGGCATCTCCCGGTACAGTGAAGTCCCTTCTGCTTCACTGCAAGTGGC 1689

QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238  
 DB 1690 TCTGGGCCGAGATTTCATCTTCCATCATCAGAGTGTGAGGCTGAAGACCTGGCAGTTAT 1749  
 QY 239 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyTyrIleGluIle 258  
 DB 1750 TTCTGTCAAGCAATTTTCGATCTCATTCACGTTGGCTCGGGGACAAAATTGGAGATC 1809  
 QY 259 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer 272  
 DB 1810 AAAGCTAGAGATCTCTCGATCGAGAGAAAGATTTTCACGC 1851  
 RESULT 9  
 ID AAT29412 standard; DNA; 1919 BP.  
 XX  
 AC AAT29412;  
 XX  
 DT 16-OCT-2003 (revised)  
 DT 19-AUG-1996 (first entry)  
 XX  
 DE DGAL4-DT-scfv(FRP5) gene in pSW50-GD5.  
 XX  
 KW Nucleic acid transfer system; gene transfer; gene therapy;  
 KW cell targeting; multidomain protein; vector; cancer; GAL4;  
 KW diphtheria toxin; single chain antibody; scFv; FRP5; ss.  
 XX  
 OS Escherichia; coli.  
 OS Saccharomyces; cerevisiae.  
 OS Corynebacterium; diphtheriae.  
 OS Mus sp.  
 OS Chimeric.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1908  
 FT sig\_peptide /\*tag= a  
 FT /\*tag= b  
 FT mat\_peptide 64..1905  
 FT /\*tag= c  
 FT /product= "GAL4-DT-scfv(FRP5) fusion protein"  
 XX  
 PN MO9613599-AI.  
 PD 09-MAY-1996.  
 PF 31-OCT-1995; 95WO-EP04270.  
 PR 01-NOV-1994; 94EP-00810627.  
 PA (WEL6/) WEL5 W.  
 XX  
 PI Wels W, Fominaya J;  
 XX  
 DR WPI; 1996-239505/24.  
 DR P-PSDB; AAR95056.  
 PT Nucleic acid transfer system for gene therapy, e.g. against cancer -  
 PT includes toxin translocation domain to target nucleic acid to specific  
 PT cell.  
 XX  
 XX Example 19; Page 78-81; 106pp; English.  
 CC Plasmid pSW50-GD5 contains a fusion gene (AAT29412) coding for a  
 CC multidomain fusion protein (AAR95056) comprising an E. coli ompa signal  
 CC peptide, a FLAG epitope, amino acids 2-147 of yeast GAL4, amino acids 196  
 CC -384 of diphtheria toxin (DT) and a single chain antibody, scFv(FRP5),  
 CC based on a monoclonal antibody raised against the HER2 antigen of human  
 CC tumour cells, all separated by spacer peptides. The GAL4 moiety acts as a  
 CC DNA binding domain, DT as a translocation domain, and scFv(FRP5) as a  
 CC ligand domain. The multidomain protein, which can be expressed in E.



CC coli, is used with an effector nucleic acid in a novel nucleic acid  
 CC transfer system suitable for gene therapy. The effector nucleic acid  
 CC includes the gene to be delivered and a cognate structure for the GAL4  
 CC DNA binding domain. (Updated on 16-Oct-2003 to standardise OS field)

XX Sequence 1919 BP; 534 A; 459 C; 462 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.19e-93	Length:	1919
Score:	1308.00	Matches:	245
Percent Similarity:	97.64%	Conservative:	3
Best Local Similarity:	96.46%	Mismatches:	0
Query Match:	53.58%	Indels:	0
DB:	2	Gaps:	0

US-09-596-774-6 (1-461) x AAT29412 (1-1919)

Qy	19	GlnlleglnleuvalglnserglyprogluleulyslypProglYgluThrVallysile	38
Db	1147	CAGGTACAACTGCAGCAGTCTGCAGCTGAATGAAAGACCTGAGAGACAGTCAAGATC	1206
Qy	39	SerCyblybAlaSerGlyTyrrProphetraentyrglymetAsntrpVallysglnala	58
Db	1207	TCCTCAAGGCTCTGGGTATCTTTCACAACTGAACTGAGTGGGTGAAGCAGGCT	1266
Qy	59	ProglYglnGlyleuYserTrpMetGlyTrpIleAsnThrSerThGlylserThrpe	78
Db	1267	CCAGGACAGGGTTTAAAGTGGATGGCTGGATTAACCTCCAGCTGAGAGACCAATT	1326
Qy	79	AlaAspAspPheLySGlyArgPheAspPheSerLeuGluThrSerAlaAsnThralaTy	98
Db	1327	GCTGATGACTTCAAGGAGCGTTTACCTCTTGGAAACCTCGCCAACTGCTGCTAT	1386
Qy	99	LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrrPheCysAlaArgTrpGlu	118
Db	1387	TTGCAGATCAACMACTCAAAAGTGAAGCATGGCTACATATTTCTGTGCAGATGGAG	1446
Qy	119	ValTYRHAGLYTYrValProfYrrTrpGlynglnYThrThValThrValSerSerGly	138
Db	1447	GTTTACCAAGGCTTACGTCCTTACTGGGGCCAAAGGACCAAGCTGCTTCTCTGGC	1506
Qy	139	GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrgln	158
Db	1507	GGTGGCGGTTCTGGTGGCGGTGCCCGCGGTGGCGTTGACATCCAGCTGACCCAG	1566
Qy	159	SerHisLysPheLeuSerThrSerValGlyAspAsgValSerIleThrCysLysAlaSer	178
Db	1567	TCTCACAAAATCTCTCCACTTCAGTAGAGACAGGGTCCAGCATCCTGCAGAGCCAGT	1626
Qy	179	GlnAspValTYrAsnAlaValAlaTrpYrrGlnGlnLysProglYglnSerProLysLeu	198
Db	1627	CAGGATGATTAATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1686
Qy	199	LeuIleTyrrSerAlaSerSerArgTyrrThrglyValProSerArgPheThrglySerGly	218
Db	1687	CTGATTTACTGGGATCTCTCCCGGATCACTGAGTCCCTTCTGCTTCACTGCGAGTGGC	1746
Qy	219	SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTy	238
Db	1747	TCTGGGCGGATTTCACTTCCATCCATGAGAGTGGAGGCTGAGAGACCTGCGACGTTAT	1806
Qy	239	PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrThylsleuGluIle	258
Db	1807	TTCTGTCAAGCAATTTCTGATCTCATTCACGTTCTGGGGGACAAATTTGAGATC	1866
Qy	259	LysAlaLeuGluIleSerAsnSerValMetTyrrPheSerSer	272
Db	1867	AAAGCTTAGAGGATCTCTCGAGTAGAGAAATTTTACGCC	1908

RESULT 10  
 AAT42038  
 ID AAT42038 standard; DNA; 2214 BP.  
 XX

AC AAT42038;  
 XX  
 XX 29-JAN-1997 (first entry)  
 DT  
 XX  
 XX Plasmid pMS238-5-TGF fragment encoding scFv(FRP5)/TGF alpha-ETA.  
 DE  
 XX  
 XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erdb-2;  
 KW epidermal growth factor; receptor; plasmid pMS238-5-TGF; cancer;  
 KW exotoxin A; ETA; transforming growth factor alpha; TGF; antitumour; ss.  
 XX  
 XX Homo; sapiens.  
 OS Mus; sp.  
 OS Pseudomonas; aeruginosa.  
 OS Synthetic.  
 OS Chimeric.  
 XX  
 XX Key Location/Qualifiers  
 FH CDS 1..2100  
 FT /tag= a  
 FT sig\_peptide 1..63  
 FT /tag= b  
 FT /note= "Ompa signal peptide"  
 FT mat\_peptide 64..2097  
 FT /tag= c  
 FT /product= "scFv(FRP5)/TGF alpha-ETA"  
 XX  
 XX EP739984-A1.  
 PD 30-OCT-1996.  
 XX  
 XX 26-APR-1995; 95EP-00106275.  
 XX  
 XX 26-APR-1995; 95EP-00106275.  
 PR (SANT-) SAN TUMORFORSCHUNGS GMBH.  
 XX  
 XX Wels W, Schmidt M, Groner B;  
 PI WPI. 1996-478748/48.  
 DR P-PSDB; AAM05138.  
 XX  
 XX Bivalent fusion proteins that bind epidermal growth factor receptor or  
 PT analogues - and comprise at least two different cell surface binding  
 PT domain(s), useful for tumour therapy.  
 XX  
 XX Example 10; Page 25-27; 52pp; English.  
 PS  
 XX A fragment (AAT42038) of bacterial expression plasmid pMS238-5-TGF codes  
 CC for scFv(FRP5)/TGF alpha-ETA (AAM05138), comprising the single-chain  
 CC binding region of murine monoclonal antibody FRP5 (specific for human  
 CC epidermal growth factor receptor erdb-2, see also AAT42034) joined to  
 CC exotoxin A from Pseudomonas aeruginosa and human transforming growth  
 CC factor (TGF) alpha (see also AAM05137). pMS238-5-TGF can be utilised in  
 CC the prodn. of a bispecific fusion protein in bacterial (esp. E. coli)  
 CC host cells. Such fusion proteins (see also AAM05139-44) are useful as  
 CC antitumour agents  
 XX  
 XX SQ Sequence 2214 BP; 431 A; 710 C; 665 G; 408 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.41e-93	Length:	2214
Score:	1308.00	Matches:	275
Percent Similarity:	71.19%	Conservative:	24
Best Local Similarity:	65.48%	Mismatches:	59
Query Match:	53.58%	Indels:	64
DB:	2	Gaps:	5

US-09-596-774-6 (1-461) x AAT42038 (1-2214)

Qy	19	GlnlleglnleuvalglnserglyprogluleulyslypProglYgluThrVallysile	38
Db	115	CAGGTACAACTGCAGCAGTCTGCAGCTGAATGAAAGACCTGAGAGACAGTCAAGATC	174

```

QY 39 SerCySLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 58
Db 175 TCCTGCAAGGCGCTCTGGGTAATCCTTTCAACAATATGAAATGAACTGGGTGAAGCAGGCTT 234
QY 59 ProGlyGlnGlyLeuLysTrpMetGlyTrpLLeuAsnThrSerThrGlyLysSerThrPhe 78
Db 235 CCAGGACAGGCTTTAAAGTGAATGAGGCGCTGATTTAAACACTCCACATGAGAGTCAACATTT 294
QY 79 AlaAspAspPheLysGlyLysPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
Db 295 GCTGATGACTTCAAGGAGCGGTTGACTTCTCTTGGAACCTCTGCAACACTCGCTAT 354
QY 99 LeuGlnLLeuAsnLeuLysSerGlyLysMetAlaThrTyrPheCysAlaArgTrpGlu 118
Db 355 TTGCAGATCAACACCTCAAAAGTGAAGCATGGCTACATATTTCTGTGCAAGATGGAG 414
QY 119 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 138
Db 415 GTTTACACGCGCTACGTTCTTACTGGGCGCAAGGAGCACCGTCAACGCTTCTCTGGC 474
QY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspLLeuLeuThrGln 158
Db 475 GGTGGCGGTTCTGTGGCGGCTGCGCTCCGCGGTTGGGTTGACATCCAGCTGACCG 534
QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerLLeuThrCysLysAlaSer 178
Db 535 TCTCACAATTCCTGCTCCACTTCAGTAGAGACAGGGTCAGATCACTCGCAAGCCAGT 594
QY 179 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 198
Db 595 CAGGATGATGAATGAATGCTGTGCTGATCAACAGAAACCAAGCAATCTCTCAACTT 654
QY 199 LeuLLeuTyrSerAlaSerSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
Db 655 CTGATTTACTGGGCACTCCCTCCGTAACACTGAGTCCCTTCTCGCTTCACTGGCAGTGGC 714
QY 219 SerGlyProAspPheThrPheThrLLeuSerSerValGlnAlaGluAspLeuAlaValTyr 238
Db 715 TCTGGGCGGATTCACCTTCAACCTGAGAGTGGCAGGCTGAGAGCTGAGAGCTGAGTAT 774
QY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnLLeu 258
Db 775 TTCTGTGCAACATTTCTGATCTCATTCACGTTGCGGCGGAGCAAAATGAGAGTGC 834
QY 259 LysAlaLeuGlnLLeuSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278
Db 835 AAAGCTTAGAGCA----- 848
QY 279 LysValaAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 298
Db 849 -----CCATCATCACTCACTCACTAGA 869
QY 299 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgLysSerValLysGlyThr 318
Db 870 GGG---CGGCAAGCTCGGCGCGCTGAC-----CGCGCACAGGCGCTGCCACCTGCCGCT 920
QY 319 GlyLeuAspPheLeuGluAspProLysLeuCysTyrLLeuAspGlyLLeuPheLLeu 338
Db 921 GGAGACTTTCAACCGTCACTCGCAGCGCGGCTGGGAGCAACTGGAGGAGCTCGGCTA 980
QY 339 TyrGlyValLLeuLLeuThrAlaLeuTyrLLeuArgAlaLysPheSerArgSerAlaGluTh 358
Db 981 TCCGCTGACGCGGCTGCTGCTCTTCACTGCGCGCGAGTGTGCA----- 1026
QY 358 TAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGly 378
Db 1027 -----TGAAACAGGTCAACA-GGT----- 1046
QY 378 uGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyLysGlnGly 398
Db 1047 -----GATCCGCAACGCGCTGCGCAGCCCGC 1072
QY 398 nArgArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 418

```

```

Db 1073 GCAAGCGCGGAGCACTGGCGCAAGCGATCCGCGACAGCCGCGACGCGCTGTGGCC 1132
QY 418 lualATyrSerGluLLeuGlyThrLysGlyGluAspArgArgGlyLysGlyHisAsp 436
Db 1133 TGACCCCTGCGCGCGCGGAGAGCGAGCGGCTTCGTCGCGGACAGGAGGACCGGACAGC 1168

RESULT 11
AAT42039
ID AAT42039 standard; DNA; 2793 BP.
XX
AC AAT42039;
XX
DT 29-JAN-1997 (first entry)
DE Plasmid pMS238-5-225 fragment encoding scFv2 (FRP5/225) -ETA.
XX
KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pMS238-5-225; cancer;
KW exotoxin A; ETA; antitumour; ss.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key
FT CDS
FT 1..2679 Location/Qualifiers
FT /*tag= a
FT sig_peptide
FT 1..63
FT /*tag= b
FT /note= "Ompa signal peptide"
FT mat_peptide
FT 64..2676
FT /*tag= c
FT /product= "scFv2 (FRP5/225) -ETA"

XX
PN EP739984-A1.
XX
PD 30-OCT-1996.
XX
PF 26-APR-1995; 95BP-00106275.
XX
PR 26-APR-1995; 95BP-00106275.
XX
PA (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
PI Wels W, Schmidt M, Groner B;
XX
DR WPI; 1996-478748/48.
XX
DR P-PSDB; AAM05139.
XX
PT Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s) - useful for tumour therapy.
XX
PS Example 11; Page 28-30; 52pp; English.
XX
CC A fragment (AAT42039) of bacterial expression plasmid pMS238-5-225 codes
CC for scFv2 (FRP5/225)-ETA (AAM05139), and is obt'd. by ligating DNA
CC fragments coding for the single-chain binding region of murine monoclonal
CC antibody FRP5 (see also AAT42034), portions of Pseudomonas aeruginosa
CC exotoxin A and the single-chain binding region of murine monoclonal
CC antibody 225 (see also AAT42033). pMS238-5-225 can be utilised in the
CC prodn. of bivalent fusion protein in bacterial (esp. E. coli) host cells.
CC Such fusion proteins (see also AAM05138-44) bind the epidermal growth
CC factor receptor and are useful as antitumour agents
XX
SQ Sequence 2793 BP; 593 A; 847 C; 806 G; 547 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,87e-93 Length: 2793
Score: 1308.00 Matches: 275
Percent Similarity: 71.19% Conservative: 24

```

Best Local Similarity: 65.48% Mismatches: 59  
 Query Match: 53.58% Indels: 64  
 DB: 2 Gaps: 5

US-09-596-774-6 (1-461) x AAT42039 (1-2793)

```

19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValIle 38
115 CAGGACGACCTGACAGCTGAGCTGAGCTGAGGAGAGAGAGAGAGAGAGAGATC 174
39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIle 58
175 TCTGTCAGAGCGCTCGGGTATCTTTCACAACTAGATGATGAGGATGAGGAGCT 234
59 ProGlyGlnGlyLeuLysTyrMetGlyTyrPheAsnThrSerThrGlyGlnSerPhe 78
235 CCAGGACAGGGGTTTAAAGTGAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 294
79 AAlaAspPheLysGlyValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
295 GCTGATGACTTCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 354
99 LeuGlnIleAsnAsnLeuLysSerGlyLysPheMetAlaThrTyrPheCysAlaArgTyr 118
355 TTGACGATCAACCACTCAAAAGTGAAGAGCATGAGCTATATTTCTGTCAGAGAG 414
119 ValTyrHleGlyTyrValProTyrTrrGlyGlnGlyThrValThrValSerGly 138
415 GTTAAACACGCGTACGTTCTTACCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 474
139 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158
475 GGTGGCGGTTCTGGTGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 534
159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178
535 TCTCACAAATTCCTGCTCACTTCAGTAGAGAGAGAGGAGGAGGAGGAGGAGGAGGAG 594
179 GlnAspValTyrAsnAlaValAlaTrrPyrGlnGlnLysProGlyGlnSerProLysLeu 198
595 CAGGATGATGAATTAAGTGTGCTGCTGATCAACAGAAACAGGAGCAATCTCTTAACCT 654
199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
655 CTGATTTACTGGGATCTCTCCGAGTACAGTGAAGTCCCTCTCGCTTCACTGGAGTGGC 714
219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238
715 TCTGGGCGGATTTCACTTTCACATCAGAGTGTGAGAGGCTGAGAGAGCTGGCAGTTAT 774
239 PheCysGlnGlnIleAspPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 258
775 TTCGTGTCAGCAATTTTCGTACTCATTCAGTTCGCTCGGAGCAAAATTGAGAGATC 834
259 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 278
835 AAAGCTCTAGAGCA----- 848
279 LysValAsnSerThrThrLysProValLeuArgThrProSerProValHisProThr 298
849 -----CCATCATCAACATCACTGAGA 869
299 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValIleGlyThr 318
870 GGG---CGGACGCGCGCGCTGAC-----CGGACACCGAGGCTGCACTGCGCT 920
319 GlyLeuAspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPheIle 338
921 GAGAGCTTTCACCCCTCATCGCCACCGCGCGCTGGAGAACACTGAGAGCACTGGCGCTA 980
339 TyrGlyValIle-IleThrAlaLeuTyrIleuArgAlaLysPheSerArgSerAlaGluTh 358
981 TCCGGTCAGGCGGCTGCTGCTTCACTGAGCGCGAGACTGTCA----- 1026

```

```

QY 358 rAlaAlaLeuLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArgI 378
DB 1027 -----TGAACACGAGTCAACA-GGT----- 1046
QY 378 uGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyGlyLysGlnI 398
DB 1047 -----GATCCGCAACGCGCTGGCCAGCCCG 1072
QY 398 nArgArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 418
DB 1073 GCAGCGCGGCGAGCTGGGCGAAGGATCCGCGAGCAGCCGAGAGAGGCGGCTGAGGCC 1132
QY 418 LuAlaTyrSerGluIleGlyThrLysGlyLysArgArgGlyLysGlyHisAsp 436
DB 1133 TGACCTGCGCGCGCGAGAGCGAGCGCTTGTCTCGGAGAGGAGCAGCGCAAGAC 1188

RESULT 12
AAT42043
ID AAT42043 standard; DNA; 2793 BP.
XX
AC AAT42043;
XX
DT 29-JAN-1997 (first entry)
XX
DE Plasmid pMS238-5-5 encoding scFv2 (FRP5/FRP5)-ETA (version 2).
XX
KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KM epidermal growth factor; receptor; plasmid pMS238-5-5; cancer;
KW exotoxin A; ETA; antitumour; ss.
XX
OS Mus; BP.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key 1..2679 Location/Qualifiers
FT CDS 1..2679
FT sig_peptide 1..63 /tag= a
FT /note= "OmpA signal peptide"
FT mat_peptide 64..2676 /tag= b
FT /product= "scFv2 (FRP5/FRP5)-ETA"
FT FT
XX
PN EP739984-A1.
XX
PD 30-OCT-1996.
XX
PF 26-APR-1995; 95EP-00106275.
XX
PR 26-APR-1995; 95EP-00106275.
XX
PA (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
PI Weis W, Schmidt M, Groner B;
XX
DR WPI; 1996-478748/48.
DR P-PSDB; AAM05143.
XX
PT Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX
PS Example 12; Page 40-42; 52pp; English.
XX
CC A fragment (AAT42043) of bacterial expression plasmid pMS238-5-5 codes
CC for scFv2 (FRP5/225)-ETA (AAM05143), and is obcd. by ligating 2 copies of
CC DNA coding for the single-chain binding region of Pseudomonas aeruginosa
CC antibody FRP5 (see also AAT42034) and portions of Pseudomonas aeruginosa
CC exotoxin A. pMS238-5-5 can be utilised in the prodn. of bivalent fusion
CC protein in bacterial (esp. E. coli) host cells. Such fusion proteins (see

```

CC also AAW05138-44) bind the epidermal growth factor receptor and are  
 CC useful as antitumour agents  
 XX  
 SQ Sequence 2793 BP; 578 A; 858 C; 812 G; 545 T; 0 U; 0 Other;

## Alignment Scores:

Pred. No.:	Length:	Score:	Matches:
1	87e-93	1308.00	275
Percent Similarity:	71.19%	Conservative:	24
Best Local Similarity:	65.48%	Mismatches:	59
Query Match:	53.58%	Indels:	64
DB:	2	Gaps:	5

US-09-596-774-6 (1-461) x AAT42043 (1-2793)

```

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuIleuValysProGlyValuThrValylsIle 38
DB 115 CAGGTACAACTGACAGAGCTTGACCTGAACTGAAAGAAAGCCGAGAGACAGCTCAAGATC 174
QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58
DB 175 TCTGTCAAGAGCCTCGGGATCTTTCACAAACATGATGAATGAATGAGGCT 234
QY 59 ProGlyGlnGlyLeuIleTyrMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 78
DB 235 CCAGACAGAGGCTTAAAGTGAAGTGGCTGATTAACACCTCCACCTGAGAGATCAACATTT 294
QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
DB 295 GCTGATGACTTCAGAGAGAGAGGCTTACCTTCTTGGAAACCTCGCAACACCTGCTAT 354
QY 99 LeuGlnIleAsnAsnLeuIleSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
DB 355 TTGCAGATCAACAACCTCAAAAGTGAAGACATGGCTACATTTCTGTGCAAGATGGAG 414
QY 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 138
DB 415 GTTACCAACGGCTACGCTTCTTACCTGAGGAGGAGGACAGGACCGCTTCTGCTGGC 474
QY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlyGlySerAspIleGlnLeuThrGln 158
DB 475 GGTGGCGGCTTCTGTGGCGGAGGCTCCCGCGGAGGCTTCTGACATCCACCTACCCAG 534
QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer 178
DB 535 TCTCAACAATTCCTGTCCACTTCAGTAGGAGACAGGGGTCAACATCCTGCAAGGCCAGT 594
QY 179 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 198
DB 595 CAGATGTGATTAAGCTGTGCTGCTGATCAACAGAAACAGAGACAATCTCCCTAAACTT 654
QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218
DB 655 CTGATTTACTCGGCATCTCCCGGTACACTGGATCTCTTCCCTTCACTCGGACAGTGGC 714
QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 238
DB 715 TCTGGGCGGATTTCACTTTCACCATCAGCAGTGTGACAGGCTGAAGACTGGGAGTTAT 774
QY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrIleValGlnIle 258
DB 775 TTCTGTGACGACAACTTTTCGACTCCATTCACGTTGGCTCGGGGACAAAATGGAGATC 834
QY 259 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 278
DB 835 AAGCTCTAGAGCA----- 848
QY 279 LysValAsnSerThrThrLysProValLeuArgThrProSerProValHisProThr 298
DB 849 -----CCATCATCACCATCACCCTAGA 869
QY 299 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThr 318
  
```

```

DB 870 GGG---CGGACGCTGGCGCGCTGAC-----CGCGACCAAGGCTCCACCTGCCGCT 920
QY 319 GlyLeuAspPheLeuGlnAspProLysValuCyThrLeuLeuAspGlyIleLeuPheIle 338
DB 921 GGAGACTTTCAACCCGTCATCGCCAGCGCGGCTGGGAACAACATGGAGACAGTGGGCTA 980
QY 339 TyrGlyValIle-IleThrAlaLeuTyrIleuArgAlaLysPheSerArgSerAlaGluTh 358
DB 981 TCCGATGACGGGCTGTGGCTCTTACTTGGCGGCGGACTGCA----- 1026
QY 358 ValAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArg 378
DB 1027 -----TGAACACAGGTGACCA-GGT----- 1046
QY 378 uGluTyrAspValLeuGlnLysValArgAlaArgAspProGluMetGlyLysGlnI 398
DB 1047 -----GATCCGCAACGCCCTGGCCGACGCCCG 1072
QY 398 nArgArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 418
DB 1073 GCACGCGCGGACCTTGGCGGAAGCCATCGCGAGACGCCGAGACAGCCGCTGCGCC 1132
QY 418 IuaLAtySerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp 436
DB 1133 TGACCTGGCGCGCGGAGAGAGAGCGCTTCGTCGCGGAGGCGACCGGCAAGAC 1188
  
```

## RESULT 13

AAT42041 ID AAT42041 standard; DNA; 3177 BP.

AC AAT42041;

DT 29-JAN-1997 (first entry)

XX Plasmid pMS240-5-225 fragment encoding scFv2 (FRP5/225)-ETA.

XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;

XX epidermal growth factor; receptor; plasmid pMS240-5-225; cancer;

XX exotoxin A; ETA; antitumour; ss.

XX Mus; sp.

XX Pseudomonas; aeruginosa.

XX Synthetic.

XX Chimeric.

XX Key

XX CDS

XX sig\_peptide

XX mat\_peptide

XX Location/Qualifiers

XX 1..3063

XX /\*tag= a

XX 1..63

XX /\*tag= b

XX /note= "OmpA signal peptide"

XX /\*tag= C

XX /product= "scFv2 (FRP5/225)-ETA"

XX EP39984-A1.

XX 30-OCT-1996.

XX 26-APR-1995; 95EP-00106275.

XX 26-APR-1995; 95EP-00106275.

XX (SANT-) SAN TUMORFORSCHUNGS GMBH.

XX Wels W, Schmidt M, Groner B;

XX MPI; 1996-478748/48.

XX P-PSDB; AAW05141.

XX Bivalent fusion proteins that bind epidermal growth factor receptor or

XX analogues - and comprise at least two different cell surface binding

XX domain(s), useful for tumour therapy.





```

FT      sig_peptide      /tag= a
FT      23..85
FT      /tag= b
FT      /label= ompa_signal_peptide
FT      89..445
FT      m1ec_RNA
FT      /tag= c
FT      /label= FRP5_heavy_chain_variable_domain
FT      446..490
FT      /tag= d
FT      /label= Linker
FT      491..814
FT      m1ec_RNA
FT      /tag= e
FT      /label= FRP5_light_chain_variable_domain
FT      815..2155
FT      mat_peptide
FT      /tag= f
FT      /label= phoa_coding_region
XX      EP502812-A1.
XX      09-SEP-1992.
XX      27-JAN-1992; 92EP-00810056.
XX      05-FEB-1991; 91EP-00810079.
XX      (CIBA ) CIBA GEIGY AG.
XX      Weis WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX      WPI; 1992-302096/37.
XX      DR P-PSDB; AAR26980.
XX      Recombinant antibodies directed to growth factor receptor C-erbB-2 - for
XX      PT diagnosing and treating tumours expressing C-erbB-2 e.g. breast or
XX      PT ovarian tumours.
XX      PS Disclosure; Page 34-40; 67pp; English.
XX      CC The sequence given encodes the single chain recombinant antibody
XX      CC designated Fv(FRP5)-phoa. The alkaline phosphatase gene (phoa) was used
XX      CC as a marker gene so that E. coli transformed with the fusion gene could
XX      CC be identified. The fusion gene was expressed in E. coli and the antibody
XX      CC was extracted. This recombinant antibody can be used for the qualitative
XX      CC and quantitative determination of c-erbB-2. This can be used for
XX      CC monitoring or in-vivo localisation of tumours overexpressing c-erbB-2.
XX      CC (Updated on 25-MAR-2003 to correct FN field.)
XX      CC
XX      SQ Sequence 2233 BP; 566 A; 573 C; 606 G; 488 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.54e-93 Length: 2233
Score: 1303.00 Matches: 249
Percent Similarity: 91.37% Conservative: 5
Best Local Similarity: 89.57% Mismatches: 9
Query Match: 53.38% Indels: 16
DB: 2 Gaps: 2

US-09-596-774-6 (1-461) x AAQ28257 (1-2233)
QY      1 MetAlATPValTTPThleuLeuPheLeuMetAlaAlaAlaValProlysGlnIle 20
DB      45 TTGCAGTGGCACTGGCTGTTCCGTACC-GTAGCGCAAGCT-----TCTCAGATA 94
QY      21 GlnLeuValGlnSerGlyProGlnLeuValysProGlyGlnThrValIleSerCys 40
DB      95 CAACCTGCAGAGCTGAGACTGAAGAAAGCTGAGAGACAGTCAGATCTCTGC 154
QY      41 LysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAlaProGly 60
DB      155 AAGGCTCTGGGTATCTTTCCAAACTATGAAATGAATGAGCTGGTGAACAGGCTCCAGGA 214
QY      61 GlnGlyLeuValTyrMetGlyTyrIleAsnThrSerThrGlyGluSerThrPheAlaAsp 80

```

```

DB      215 CAGGGTTTAAAGTAGGAGGCTGGATTAAACCTCCAGCTGAGAGTCAACATTGGCTGAT 274
QY      81 AspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyrLeuGln 100
DB      275 GACTTCAAAGGACGGTTTGACTTCTCTTGGAAACCTTGCCAAACACTGCTATTGGCAG 334
QY      101 IleAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGlnValTyr 120
DB      335 ATCAACAACCTCAAAAGAGAAAGACATGGCTCATATTTCTTGCAAGATGGAGGTTTAC 394
QY      121 HisGlyTyrValProTyrTTPGlyGlnGlyThrThrValThrValSerSerGlyGlyGly 140
DB      395 CACGGCTACGTTCTTACTGGGGCCAAAGGAGCAACGGTCACCGTTTCTCTGGCGGTGC 454
QY      141 GlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGlnSerHis 160
DB      455 GGTTCGTGGCGGGGTGGCTGGCGGGCGGTGGCTGTGACATCACTGACCTACAGTCTAC 514
QY      161 LysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSerGlnAsp 180
DB      515 AAATTCCCTGTCACCTTCAGTAGAGACAGGCTCAGCATCAGTCAGAGCCAGTCAGAT 574
QY      181 ValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeuLeuIle 200
DB      575 GTGTATATGCTGTGCTGGTATCAACAGAAACAGAGCAATCTCTAAACTTCTGATT 634
QY      201 TyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 220
DB      635 TACTGGGATCTCCCGGTACACTGAGAGTCCCTTCTGCTTCACTGGAGTGGCTTGG 694
QY      221 ProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCys 240
DB      695 CCGGATTTCACTTTCACCATATGACAGTGTGACAGCTGAAGACTGGCAGTTATTTCGT 754
QY      241 GlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIleLysAla 260
DB      755 CAGCAACATTTTCGTACTCCATTCACGTTCCGCTCGGGGACAAATGAGATCAAAAGCT 814
QY      261 LenglIleSerAsnSerValMetCysPheSerSerValValProValLeuGln 278
DB      815 CTAGAG-----CCGTGTTCTGGA 832

```

Search completed: April 20, 2005, 10:44:24  
 Job time : 767.356 secs

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 20, 2005, 10:19:55 ; Search time 788.902 Seconds

(without alignments)  
3545.124 Million cell updates/sec

Title: US-09-596-774-6

Perfect score: 2441

Sequence: 1 MAMWTLFLMAAKVPRQI.....LSTATKDTYDALHMQTLAPR 461

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303335566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODL=frame\_plus\_p2n.model -DEV=xlh  
-O=/cg2\_1/USPTO\_spoor1\_h/US09596774/runatc\_20042005\_075444\_12560/app\_query.faste\_1.1230  
-DB=Published Applications NA -OFMT=factap -SUFFIX=apr20.rnpb -MINMATCH=0.1  
-LOOPEL=0 -LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blousum62  
-TRAUS=human40.cdi -LIST=45 -DOCLIGN=200 -THR SCORE=spc -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAXLEN=200000000 -USER=US09596774 -CGEN\_1\_1\_598 -ORUNATC\_20042005\_075444\_12560  
-NCPU=6 -ICPU=3 -NO\_MMAP -LARGEQUERY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

- 1: /cg2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cg2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cg2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cg2\_6/ptodata/1/pubpna/US06\_PUBCOMB.seq:\*
- 5: /cg2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cg2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cg2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cg2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cg2\_6/ptodata/1/pubpna/US09A\_PUBCOMB.seq:\*
- 10: /cg2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cg2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cg2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cg2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cg2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 15: /cg2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 16: /cg2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 17: /cg2\_6/ptodata/1/pubpna/US10G\_PUBCOMB.seq:\*
- 18: /cg2\_6/ptodata/1/pubpna/US10H\_PUBCOMB.seq:\*
- 19: /cg2\_6/ptodata/1/pubpna/US10I\_NEW\_PUB.seq:\*
- 20: /cg2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cg2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cg2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1249	51.2	1920	US-10-120-198B-1	Sequence 1, Appli
2	1249	51.2	6834	US-10-120-198B-5	Sequence 5, Appli
3	1015.5	41.6	7654	US-10-006-773-1	Sequence 1, Appli
4	1015.5	41.6	7654	US-10-006-771A-1	Sequence 1, Appli
5	994	40.7	771	US-10-013-173-38	Sequence 38, Appli
6	994	40.7	771	US-10-150-762-38	Sequence 38, Appli
7	994	40.7	771	US-10-244-821-38	Sequence 38, Appli
8	981.5	40.2	909	US-09-887-853-1	Sequence 1, Appli
9	969.5	39.7	909	US-09-883-547-1	Sequence 1, Appli
10	930	38.1	1350	US-08-812-393A-1	Sequence 1, Appli
11	928	38.0	1350	US-09-774-681-1	Sequence 1, Appli
12	917.5	37.6	1515	US-10-239-656-78	Sequence 78, Appli
13	913	37.4	816	US-09-766-543-9	Sequence 9, Appli
14	903.5	37.0	723	US-10-127-890-90	Sequence 90, Appli
15	903.5	37.0	723	US-10-717-243-90	Sequence 90, Appli
16	901.5	36.9	804	US-10-071-485-1	Sequence 1, Appli
17	897	36.7	840	US-09-766-543-11	Sequence 11, Appli
18	888.5	36.4	1626	US-10-071-485-84	Sequence 84, Appli
19	888.5	36.4	2133	US-10-071-485-89	Sequence 89, Appli
20	886.5	36.3	1509	US-10-239-656-74	Sequence 74, Appli
21	875.5	35.9	10511	US-10-059-261-109	Sequence 109, Appli
22	875.5	35.9	10511	US-10-627-649-109	Sequence 109, Appli
23	872.5	35.7	729	US-10-879-994-9	Sequence 9, Appli
24	872.5	35.7	729	US-10-610-452-9	Sequence 9, Appli
25	862	35.3	1869	US-10-378-832A-1	Sequence 1, Appli
26	860.5	35.0	761	US-10-354-246-4	Sequence 4, Appli
27	854.5	35.0	1509	US-10-239-656-76	Sequence 76, Appli
28	854	35.0	777	US-10-096-246-5	Sequence 5, Appli
29	854	35.0	777	US-10-096-246-7	Sequence 7, Appli
30	848.5	34.8	1497	US-10-239-656-72	Sequence 72, Appli
31	848	34.7	702	US-10-071-485-92	Sequence 92, Appli
32	846	34.7	777	US-10-096-246-6	Sequence 6, Appli
33	841.5	34.5	756	US-09-984-186-17	Sequence 17, Appli
34	841.5	34.5	756	US-10-237-667-17	Sequence 17, Appli
35	841.5	34.5	756	US-10-237-667-17	Sequence 17, Appli
36	841.5	34.5	756	US-10-237-866-17	Sequence 17, Appli
37	841.5	34.5	756	US-10-237-871-17	Sequence 17, Appli
38	841.5	34.5	756	US-10-237-624-17	Sequence 17, Appli
39	841.5	34.5	756	US-10-702-536-17	Sequence 17, Appli
40	841.5	34.5	756	US-10-702-536-17	Sequence 17, Appli
41	839	34.4	921	US-10-363-349-12	Sequence 12, Appli
42	838	34.3	822	US-09-749-873-108	Sequence 108, Appli
43	833.5	34.1	2090	US-10-104-522-6	Sequence 6, Appli
44	833.5	34.1	2090	US-10-060-585-6	Sequence 6, Appli
45	833.5	34.1	2090	US-10-334-235-6	Sequence 6, Appli

## ALIGNMENTS

RESULT 1  
US-10-120-198B-1  
Sequence 1, Application US/10120198B  
Publication No. US20030215427A1  
GENERAL INFORMATION:  
APPLICANT: Jensen, Michael  
TITLE OF INVENTION: CEF-SPECIFIC REDIRECTED IMMUNE CELLS  
FILE REFERENCE: 1954-337  
CURRENT APPLICATION NUMBER: US/10/120, 198B  
PRIOR APPLICATION NUMBER: 60/282, 859  
PRIOR FILING DATE: 2001-04-11  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1920  
TYPE: DNA  
ORGANISM: artificial sequence  
FEATURE:  
OTHER INFORMATION: mouse-human chimera

```

FEATURE:
; NAME/KEY: CDS
; LOCATION: (14) (1906)
; OTHER INFORMATION: scfvc construct
US-10-120-198B-1

Alignment Scores:
Pred. No.: 8,64e-121 Length: 1920
Score: 1249.00 Matches: 276
Percent Similarity: 54.60% Conservative: 62
Best Local Similarity: 44.59% Mismatches: 96
Query Match: 51.17% Indels: 186
DB: 17 Gaps: 12

US-09-596-774-6 (1-461) x US-10-120-198B-1 (1-1920)

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 38
DB 80 CAGGTCCAACTGCACAGCTGGGCTGMACTGTGAAGCCCTGGGGCTTCACTGAAGCTG 139
QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysIle 58
DB 140 TCCTCAAGGCTTCTGGCTACACCTTCCACCGCTACTGAGTCACTGGGTGAAGCAGG 199
QY 59 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 78
DB 200 CCGTGAACATGGCTTGAAGTGAATGAGATTAATCTCAGCAAGCGTCTGACTACTAC 259
QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98
DB 260 AATGAGAGTTTCMAAGACAGACCCACTGACTGTAGCAAAATCTCCACACAGCCCTTC 319
QY 99 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 118
DB 320 ATGCACACTCAGCGGCTGACATCTGAGACTCTGCAAGCTTATTTCTGTGCAAGA----- 373
QY 119 ValTyrHisGly-----TyrValProTyrTyrGlyGlnGlyTyrThrValThrVal 135
DB 374 GATTACTACGGTACTAGCTACAACTTGTGACTACGCGGCAAGGACCACTCTCACAGTC 433
QY 136 SerSerGlyGlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerAspIleGln 155
DB 434 TCCCTGAGAGGTGGGTGAGTGGAGGTGGCGATCCGGTGGCGAGGTGAGTGCATTCAG 493
QY 156 LeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCys 175
DB 494 ATGACACATCTTCATCTCTCTTTCTGTATCTCTAGAGACAGAGTCACTTACTTGC 553
QY 176 LysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSer 195
DB 554 AAGGCTATGAAAGCATTAATAATCGGTTCCTGTATCAGACAGACACCGAATAATCT 613
QY 196 ProLysLeuLeuLysTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThr 215
DB 614 CTTAGGCTCTTAATATCTGTGTCACCAATTTGTACTGGGGTTCCTTCAAGATTAGT 673
QY 216 GlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeu 235
DB 674 GCGAGTGATCTGAAAGGATTACACTCTCACATTCACAGTCTTCAGGCTGAAGATTTT 733
QY 236 AlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrLys 255
DB 734 GCTACTTAATATCTGTCAACATATTGAGTACTCATTCAGTTCCGCTCGGGGACAGAG 793
QY 256 LeuGluLys----- 259
DB 794 CTGAGATCAAAAGTAGAACCAAAATCTTTCGACAAAACCTACACATGCCACCTGCCCA 853
QY 259 ----- 259
DB 854 GCACTGAAGTCTGGGGGAGCGTCACTTCTCTTCCCCCAAAACCAAGAGACACC 913
QY 260 -----AlaLeuGluLysSer----- 264

```

```

DB 914 CTATATCTCCCGAGACCCCTGAGTCAATGCTGTGTGTGAGCTGAGCCAGAAAC 973
QY 264 ----- 264
DB 974 CCGAGGTCAAGTTCAACTGTGTACGTGACGCGGTGAGGTGATTAATGCCAAGAAAC 1033
QY 265 -----AsnSerValMetTyrPheSerSerValValProValLeuGln 1093
DB 1034 CCGCGGAGAGAGAGTACAACACAGCAGTACCGTGTGTGTCAGCGTCTCACCGCTTCGAC 1093
QY 279 -----LysValAsnSerThrThrThrLysPyr 287
DB 1094 CAGAGCTGCTGATGCGAAGAGTACAAAGTCAAGGTCTTCCAAAGCCCTCCAGCC 1153
QY 287 OValLeuArgThrProSerProValHisProThrGlyTyrSerGlnProGlnArgProG 307
DB 1154 CCGATCGAAGAAACATCTCCA--AAGTCAAAGGCAAGCCCGAAGAACACAGGTGTAC 1210
QY 307 uAspCys----- 309
DB 1211 ACCCTGCACCATCAAGATGAGCTGACCAAGAACAGGTCACTGACTGCTGTGTC 1270
QY 310 -----ArgProArgGlySerValLysGly 317
DB 1271 AAAGGCTTATATCCAGCAGCATGCGCGTGTGAGTGGAGAGCATGGCGACCGGAGAAC 1330
QY 317 YThrGlyLeuAspPheLeuGluAspProLysLeu--CysTyrLeuLeuAspGlyIleLe 336
DB 1331 AACCT-----ACAAGACCAAGCTTCGCTGTGATCC--GACGGCTCTTT 1374
QY 336 uPheIleTyr----- 339
DB 1375 CTTCTCTACAGCAAGCTCACCGTGACAAAGAGGTGGACAGAGGAAAGCTTTCTC 1434
QY 339 ----- 339
DB 1435 ATGCTCGTATGATGAGGCTGTGCACAAACACTACAGCAGAAAGCCTCTCCCTGTC 1494
QY 340 -----GlyValIleIle----- 343
DB 1495 TCCCGGAAATATGCCCTGATTTGTGTGAGGGCGTGGCGGCTCTCTTTCATTGG 1554
QY 344 -ThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThrAlaAlaAsnLeuG 363
DB 1555 GCTAGGCAATTTCTTACAGTGAAGTTTCAAGAGAGCCAGACGCCCCCGCTACAGCA 1614
QY 363 nAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArgGluGluTyrAspValLe 383
DB 1615 GGGCCAGAAACCACTCTATTAAGAGCTCAATCTAGAGCAAGAGAGAGTACGATGTTT 1674
QY 383 uGluLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGlnArgArgAsnPr 403
DB 1675 GACAAAGAGACGTGGCGGACCTTGAATGGGGGAAAG--CCGAGAGAGAAAGAACCC 1731
QY 403 oGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSerGluI 423
DB 1732 TCAGAAAGCCCTGTACAAAGAACTGACAGAAAGTAAAGTGAAGTGGGAGGCTTACAGTGA 1791
QY 423 eGlyThrLysGlyGluArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSe 443
DB 1792 TGGAGTGAAGAGGAGCGCGCGGAGGGGCAAGGGGACAGTGTGCTTACAGAGGTCTG 1851
QY 443 rThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThrLysAlaProArg 461
DB 1852 TACAGCCACCAAGAGACCTTACAGAGCTTCAATGACAGGCTTGGCCCTTCGC 1906

```

## RESULT 2

```

US-10-120-198B-5
; Sequence 5, Application US/10120198B
; Publication No. US20030215427A1
; GENERAL INFORMATION:
; APPLICANT: Jensen, Michael

```

TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS  
FILE REFERENCE: 1954-337  
CURRENT APPLICATION NUMBER: US/10/120,198B  
CURRENT FILING DATE: 2002-04-11  
PRIOR APPLICATION NUMBER: 60/282,859  
PRIOR FILING DATE: 2001-04-11  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 5  
LENGTH: 6834  
TYPE: DNA  
ORGANISM: artificial sequence  
FEATURE:  
OTHER INFORMATION: plasmid construct  
US-10-120-198B-5

## Alignment Scores:

Pred. No.:	4,87e-120	Length:	6834
Score:	1249.00	Matches:	276
Percent Similarity:	54.60%	Conservative:	62
Best Local Similarity:	44.59%	Mismatches:	96
Query Match:	51.17%	Indels:	186
	17	Gaps:	12

US-09-596-774-6 (1-461) x US-10-120-198B-5 (1-6834)

QY 19 GlnIleGlnLeuValGlnSerGlyProGlyLeuValGlySerProGlyValThrValIle 38  
DB 1878 CAGGTCACACGACGACGCTGGGCTGAACCTGGAAGGCTTCACTGAGTGAAGCTG 1937  
QY 39 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 58  
DB 1938 TCCTGCAAGGCTTCTGCTACACCTTCAACGGCTACCTGATGCACTGAGTGAAGAGAG 1997  
QY 59 ProGlyGlnGlyLeuValTyrPheGlyTyrIleAsnThrSerThrGlyGlySerThrPhe 78  
DB 1998 CCTGACACATGCTGATGATGATGAGAGATTAACTCTAGCAAGGCTGCTACTACTAC 2057  
QY 79 AlaAspAspPheGlyValArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 98  
DB 2058 AATAGAGAGGTCGAAGAGCAAGGACCACTGACTGAGCAAACTCTCAACGACCTTTC 2117  
QY 99 LeuGlnIleAsnLeuValSerGlyIleAspMetAlaThrTyrPheCysAlaArgTyrPhe 118  
DB 2118 ATGCACTACACGCGCTCACTCTGAGACTCTGCACTATTTCTGCAAGA----- 2171  
QY 119 ValTyrHisGly-----TyrValProTyrTyrGlyGlnGlyTyrThrValThrVal 135  
DB 2172 GATTACTACGCTAGCTAGCTACCACTTGTACTGAGGCGCAAGGACCACTCTCACTG 2231  
QY 136 SerSerGlyGlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGln 155  
DB 2232 TCCTCAGAGAGGCGGAGTGAAGGAGGCGGAGTCCGGTGGCGGAGTGAAGCTCAG 2291  
QY 156 LeuThrGlnSerHisGlyPheLeuSerThrSerValGlyAspArgValSerIleThrCys 175  
DB 2292 ATGACACATCTTCACTCTCTTCTGTATCTCTAGAGACAGAGTACCACTTACTTGC 2351  
QY 176 LysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSer 195  
DB 2352 AAGGCTAATGAAGCATTAATTAATCGGTACTGATCAGACAGACACCGAATAATCT 2411  
QY 196 ProLysLeuLeuIleTyrSerHisSerSerArgTyrThrGlyValProSerArgPheThr 215  
DB 2412 CTTAGGCTCTTAATATCTGTCGCAACAAATTGTGTAAGTCTTCAAGATTCACT 2471  
QY 216 GlySerGlySerGlyProAspPheThrThrIleSerSerValGlnAlaGlyAspLeu 235  
DB 2472 GCGAGTGAATCTGGAAGAGATTACCTCACTCACTTACAGCTCTTCAAGGCTGAAGATT 2531  
QY 236 AlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLys 255  
DB 2532 GCTACTATTACTGTCAACATATTGAGTCACTCATTCAGCTTGGGCTGGGAGCAGAG 2591

QY 256 LeuGlnIleLys----- 259  
DB 2592 CTCGAGATCAAAAGTAGAACCCAAATCTTCTGACAAAACCTACACATGCCACCGGCCA 2651  
QY 259 ----- 259  
DB 2652 GCACCTGAATCTTGGGGGAGACCGTCACTTCTTCTTCCCCCAAAACCAAGAGACAC 2711  
QY 260 -----AlaLeuGlnIleSer----- 264  
DB 2712 CTGATGATCTCCGAGACCCCTGAGTCAATGCGTGTGTGAGAGTGAACGACGAGAGAC 2771  
QY 264 ----- 264  
DB 2772 CCTGAGTCAAGTTCACCTGCTGAGACGCGCTGAGAGTGCATTAATGCCAAGCAAG 2831  
QY 265 -----AsnSerValMetTyrPheSerSerValProValLeuGln 278  
DB 2832 CCGGAGAGAGACAGTACACAGACGTAACCGTGTGTGCTCACTCACTGCTGAC 2891  
QY 279 -----LysValAsnSerThrThr-ThrLysPr 287  
DB 2892 CAGACTGGCTGAATGGCAAGAGTACAAAGTGCAGAGTCTCCAAAGCCCTCCAGCC 2951  
QY 287 oValLeuArgThrProSerProValHisProThrGlyTyrSerGlnProGlnArgProGly 307  
DB 2952 CCATTCGAGAAACATCTTCCA--AGCCAAAGGCGAGCCCGGAGAACACACAGGTGAC 3008  
QY 307 uAspCys----- 309  
DB 3009 ACCCTGCACATCAGACGATGAGGTGACCAAGAACAGTCAAGCTGACCTGCTGCTG 3068  
QY 310 -----ArgProArgGlySerValLysGly 317  
DB 3069 AAAGCTTCTATCCAGAGCATGCGCTGAGTGGGAGACAAATGGCGACCGGAGAAC 3128  
QY 317 YThrGlyLeuAspPheLeuGlnuAspProLysLeu---CysTyrLeuLeuAspGlyIleLe 336  
DB 3129 AACT-----ACAAAGCACGCTCCCTGCTGAGCTC-GACGGCTCTT 3172  
QY 336 uPheIleTyr----- 339  
DB 3173 CTTCCTTACAGCAAGCTCACCGTGAACAAGAGAGTGGCAGCAGGGGAAGCTTCTTC 3232  
QY 339 ----- 339  
DB 3233 ATGCTCCGTATGATGAGGCTCTGCAACACCTACACGCAAGAGCTTCTCCTGTC 3292  
QY 340 -----GlyValIleIle----- 343  
DB 3293 TCCCGGGAATGCGCCCTGATTGCTGGGGGGGCTGCCGCGCTCTTCTTTCATTGG 3352  
QY 344 -ThrIleLeuTyrLeuArgAlaLysPheSerArgSerAlaGlnThrAlaIleAsnLeuGly 363  
DB 3353 GCTAGGATCTTCTTCAAGAGTGAAGTTACAGAGAGGCGAAGCCCGCGGTACACAGCA 3412  
QY 363 nAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyValArgGlnGlnTyrAspValLe 383  
DB 3413 GGGCAGAACCCAGCTCTATTAACGACTCAATCTGAGCGAAGAGAGGTCGATGTTT 3472  
QY 383 uGluLysLysArgAlaArgAspProGlyLysGlnGlnGlnArgArgAsnPr 403  
DB 3473 GGAACAGAGAGCTGGCGGAGCCCTGAGATGGGGGAGAG---CCGAGAGAGAGAGACC 3529  
QY 403 oGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSerGlyIle 423  
DB 3530 TCAGAGAGGCTGTACATGAATGACGAGAAAGATTAATGAGGCTGACATGAGAT 3589  
QY 423 eGlyThrLysGlyLysArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSe 443  
DB 3590 TGGGATGAAGAGCGAGCGCGAGGGGCAAGGGGCGACGATGCTTTAACGAGGCTCTCAG 3649

QY 443 rThra1aThrLyseAspThrTyraSpAlaLeuH1sMetGlnThrLeuAlaProArg 461  
Db 3650 TACAGCCACCAGACACCTACACAGCCCTTACATGACGAGCCCTCCCTCGC 3704

RESULT 3  
US-10-006-773-1  
; Sequence 1, Application US/10006773  
; Publication No. US20020132983A1  
; GENERAL INFORMATION:  
; APPLICANT: Junghans, Richard P.  
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti-  
; FILE REFERENCE: 003  
; CURRENT APPLICATION NUMBER: US/10/006,773  
; PRIOR FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: 60/250,089  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 7654  
; TYPE: DNA  
; ORGANISM: Homo sapiens and Mus sp.  
; NAME/KEY: CDS  
; LOCATION: (2428)..(3759)  
; OTHER INFORMATION: Chimeric IgTcR sequence contained in retroviral vector. Retrovi-  
; OTHER INFORMATION: al vector sequence (non-coding regions) are incidental to the inv-  
; OTHER INFORMATION: ention. The translated (coding region) is relevant to the invent-  
; US-10-006-773-1

Alignment Scores:  
Pred. No.: 1,996-95 Length: 7654  
Score: 1015.50 Matches: 225  
Percent Similarity: 57.74% Conservative: 25  
Best Local Similarity: 51.96% Mismatches: 58  
Query Match: 41.60% Indels: 125  
Gaps: 8

US-09-596-774-6 (1-461) x US-10-006-773-1 (1-7654)

QY 149 GlyGlyGlySerAspAlaGlnLeuThrGlnSerHisLysPheLeuSerThrSerValGly 168  
Db 2473 GGTGCCCTCCGACATCAGTACAGCCAGAGCCAGCCAGCGGCGTGGT 2532

QY 169 AsparGlySerIleThrCysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyr 188  
Db 2533 GACAGAGTACCATCAGCTGTAAAGCCAGTACAGATGTGGTACTTCTGTAGCTTGATC 2592

QY 189 GlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThr 208  
Db 2593 CAGCAGAGCCAGGTAAAGCTCCAAAGCTGCTGATCTTACATCGACATCCAGCCGACACT 2652

QY 209 GlyValProSerArgPheThrGlySerGlySerGlyProAspPheThrPheThrIleSer 228  
Db 2653 GGTGTGCCAGACATTCAGGGGTAGCGGTAGCGGTACCACTTACCTTACCATCAGC 2712

QY 229 SerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGlnHis-----PheArgThr 246  
Db 2713 AGCTTCAGCCAGAGACATCGCCACTTACTGACAGCAATATAGCTCTATCGG--- 2769

QY 247 ProPheThrPheGlySerGlyThrLysLeuGlnIleLys----- 259  
Db 2770 -----TCGTTCCGCCCAAGGAGCCAGGTGAATCAAAAGAGGTGGCTCAGATCGGGT 2823

QY 259 ----- 259

Db 2824 GGATCCGGCTCTGTGTGCTCAGATCGAGATCGCACTGGTGAGAGCGGTGAGGTGT 2883

QY 260 -----AlaLeuGlnIleSerAsnSerVal----- 267

Db 2884 GTGCAACCTGGCGGCTCCCTGCTGCTCTGCTCGCATCTGCTGATTTTCACACA 2943

QY 267 ----- 267  
Db 2944 TATTGATGATGTTGGTGAGACAGGACCTGGAAAGGTCTTGATGTGATTTGAAATTT 3003

QY 267 ----- 267  
Db 3004 CATCCAGATGACAGTACGATTAACTATCCGCCCTCTCTAAAGATAGATTACAAATATCG 3063

QY 267 ----- 267  
Db 3064 CGAGACAAGCCAGAAACACATTGTCTTCCGAATGAGACCTGAGACCCGAAACACC 3123

QY 268 -----MetTyrPheSerSer 272  
Db 3124 GGGGTCTATTGTTGTGACACCTTTACTTCGGCTTCCCTGTTGCTTATTTGGGGCCAA 3183

QY 273 ValValProValLeuGlnLysValAsnSerThrThrLysProValLeuAlaGlyThrPro 292  
Db 3184 GGGACCCCGGTACCGTCTTCAAGTCTTAAGCCCAACAGCGCCAGCGCCGACCA 3243

QY 293 SerProValHisProThrGlyThrSerGlnPro-----GlnArgProGlnAspCysArg 310  
Db 3244 ACAACGGCG-CCACCATCGGTGCGAGCCCTGTCCCTGGCCAGAGGGCGGTCCG 3300

QY 311 ProArg-----GlySerValLysGlyThrGlyLysAspPheLeuGlnAspProLysLeu 328  
Db 3301 CAGCCGGCGGGGGCGAGTGCACAGAGAGGGGCTGAGCTTCCCTCGATCCCAAACTC 3360

QY 329 CysTyrLeuLeuAspGlyLysLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeu 348  
Db 3361 TGCTACTGCTGTGATGAGATTCCTTCACTATGATGATGATCTCAGCTCTGTTTCTCG 3420

QY 349 ArgAlaLysPheSerLysSerAlaGlnTyrAlaAlaAsnLeuGlnAspProAsnGlnLeu 368  
Db 3421 AGAGTAAAGTTGACGAGAGCGGAGCCCGCGCTTACACAGAGCGCAGAACAGCTC 3480

QY 369 TyrAsnGlnLeuAsnLeuGlnArgArgGlnGlnTyrAspValLeuGlnLysArgAla 388  
Db 3481 TATTAAGACTCAATTTAGACGAAAGAGAGAGTGTGAGTGTGACAAAGAGCTGGC 3540

QY 389 ArgAspProGlnMetGlyGlyLysGlnGlnArgArgAspProGlnGlnGlyValTyr 408  
Db 3541 CGGACCTCGAATGGGGGGAAG--CCAGAGAAGAAAGACCTCAGAGAGCTGTAC 3597

QY 409 AsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSerGlnIleGlyThrLysGlyGlu 428  
Db 3598 AATGAACGCAAGAAAGATAGATGGCGGAGGCTTACAGTGAATGGATGAGTGAAGAGCGAG 3657

QY 429 ArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLysSerThrAlaThrLysAsp 448  
Db 3658 CCGCGAGGGGCAAGAGGGGAGATGGCTTTTACAGGGTCTCAGTACAGCCAGAGC 3717

QY 449 ThrTyrAspAlaLeuHisMetGlnThrLeuAlaProArg 461  
Db 3718 ACTTACGAGCGCTTACATGACAGCCCTGCCCTCCG 3756

RESULT 4  
US-10-006-771A-1  
; Sequence 1, Application US/10006771A  
; Publication No. US20020165360A1  
; GENERAL INFORMATION:  
; APPLICANT: Junghans, Richard P.  
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen  
; FILE REFERENCE: 002  
; CURRENT APPLICATION NUMBER: US/10/006,771A  
; PRIOR FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: 60/250,090  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 7654  
; TYPE: DNA



```

Db      1 CAGGTGAAGCTGCGACGAGTCCGAGTTCAAGAACCCGGTGAAGCCCTCAAGATC
Qy      39 SerCysAlaSerGlyTyrProPheThrAspTyrGlyMetAsnTyrValysGlnAla
Db      61 AGCTGCAAGGCTTCTGTATACCTTACCGCTTGTGTATGATCACTGGGTGAAGAGGCT
Qy      59 ProGlyGlnGlyLeuYstrPheMetGlyTyrPheAsnThrSerThrGlyGluSerThrPhe
Db      121 CCGGCGAAGGCTTAAAGTGAATGGCTGGATTAACCAAACTGTGAAGCAACTAT
Qy      79 AlaAspAspPheGlyValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr
Db      181 GTTGAAGAGTTAAAGGCTGCTTTCCTTCTTGGAGACTCTGCCACCACTGCTTAT
Qy      99 LeuGlnIleAsnAlaSerGlyLeuAspPheAlaThrTyrPheCysAlaArgTyrGlu
Db      241 TTGCAGATCAACAACTCAAAATAGAGACAGCGCTAAATATTTCTGTGACGTTGGAGC
Qy      119 ValTyrHisGlyTyrVal-----ProTyrTyrGlnGlyThrThrValThrVal
Db      301 TTCTAT--GATTACGTGAAGATGATGATTAAGGCGCAAGGAGCAAGCTCACCGTC
Qy      136 Ser-----SerGlyGlyGlySerGly
Db      358 TCCAAGATCTGTGGTGGCGGCTCGCGCGGTGGGTGGGTGGGTGGGTGGGTGGGT
Qy      144 GlyGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly
Db      418 GGTGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT
Qy      163 LeuSerThrSerValGlyAspArgValSerIleThrCysAlaSerGlnAspValTyr
Db      478 ATGCCACTTCAGTAAAGTATCTGTCAAGCTGATCTGCAAGGCAAGCTGATGATGAT
Qy      183 AsnAlaValAlaTyrPyrGlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSer
Db      538 ACCAATGTGCTCGTATCAACAGAAACCGGGTCAATCCCGAAGCACTGATTTACTCG
Qy      203 AlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGlyProAsp
Db      598 GCATCTTACCGTATCAAGTGTGTCCCGGATCCGCTTACCGGAGTGTCTGGAGCCGAT
Qy      223 PheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyrPheCysGlnGln
Db      658 TTCACGCTCAACATCAGCAATGTACAGTGAAGCTTGGCGAGTATTTCTGTCACTCA
Qy      243 HisPheArgThrPro---PheThrPheGlySerGlyThrLysLeuGlnIleLys
Db      718 TATTACACTATCCGTTATTCACGTTCCGCTCGGGAGCAAGTTGAATGAAG 771

```

```

; ORGANISM: Artificial Sequence
; FEATURES:
; OTHER INFORMATION: FRA3 single chain antibody-genomic Streptavidin
; US-10-150-762-38

```

```

Alignment Scores:
Pred. No.: 1.59e-94
Score: 994.00
Percent Similarity: 82.95%
Best Local Similarity: 74.81%
Query Match: 40.72%
DB: 15
Gaps: 5

```

US-09-596-774-6 (1-461) x US-10-150-762-38 (1-771)

```

Qy      19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle
Db      1 CAGGTGAAGCTGCGACGAGTCCGAGTTCAAGAACCCGGTGAAGCCCTCAAGATC
Qy      39 SerCysAlaSerGlyTyrProPheThrAspTyrGlyMetAsnTyrValysGlnAla
Db      61 AGCTGCAAGGCTTCTGTATACCTTACCGCTTGTGTATGATCACTGGGTGAAGAGGCT
Qy      59 ProGlyGlnGlyLeuYstrPheMetGlyTyrPheAsnThrSerThrGlyGluSerThrPhe
Db      121 CCGGCGAAGGCTTAAAGTGAATGGCTGGATTAACCAAACTGTGAAGCAACTAT
Qy      79 AlaAspAspPheGlyValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr
Db      181 GTTGAAGAGTTAAAGGCTGCTTTCCTTCTTGGAGACTCTGCCACCACTGCTTAT
Qy      99 LeuGlnIleAsnAlaSerGlyLeuAspPheAlaThrTyrPheCysAlaArgTyrGlu
Db      241 TTGCAGATCAACAACTCAAAATAGAGACAGCGCTAAATATTTCTGTGACGTTGGAGC
Qy      119 ValTyrHisGlyTyrVal-----ProTyrTyrGlnGlyThrThrValThrVal
Db      301 TTCTAT--GATTACGTGAAGATGATGATTAAGGCGCAAGGAGCAAGCTCACCGTC
Qy      136 Ser-----SerGlyGlyGlySerGly
Db      358 TCCAAGATCTGTGGTGGCGGCTCGCGCGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT
Qy      144 GlyGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly
Db      418 GGTGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT
Qy      163 LeuSerThrSerValGlyAspArgValSerIleThrCysAlaSerGlnAspValTyr
Db      478 ATGCCACTTCAGTAAAGTATCTGTCAAGCTGATCTGCAAGGCAAGCTGATGATGAT
Qy      183 AsnAlaValAlaTyrPyrGlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSer
Db      538 ACCAATGTGCTCGTATCAACAGAAACCGGGTCAATCCCGAAGCACTGATTTACTCG
Qy      203 AlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGlyProAsp
Db      598 GCATCTTACCGTATCAAGTGTGTCCCGGATCCGCTTACCGGAGTGTCTGGAGCCGAT
Qy      223 PheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyrPheCysGlnGln
Db      658 TTCACGCTCAACATCAGCAATGTACAGTGAAGCTTGGCGAGTATTTCTGTCACTCA
Qy      243 HisPheArgThrPro---PheThrPheGlySerGlyThrLysLeuGlnIleLys
Db      718 TATTACACTATCCGTTATTCACGTTCCGCTCGGGAGCAAGTTGAATGAAG 771

```

RESULT 7

```

; Sequence 38, Application US/10244821
; Publication No. US20030143233A1
; GENERAL INFORMATION:

```

APPLICANT: Goshorn, Stephen Charles  
 APPLICANT: Graves, Scott Stoll  
 APPLICANT: Schultz, Joanne Elaine  
 APPLICANT: Lin, Yuhang  
 APPLICANT: Sanderson, James Allen  
 APPLICANT: Reno, John M.  
 APPLICANT: Dearstyne, Erica A.  
 TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
 TITLE OF INVENTION: METHODS OF USE THEREOF  
 FILE REFERENCE: 690022.547C3  
 CURRENT APPLICATION NUMBER: US/10/244.821  
 NUMBER OF SEQ ID NOS: 92  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 38  
 LENGTH: 771  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: PRLA3 single chain antibody-genomic Streptavidin  
 OTHER INFORMATION: fusion construct  
 US-10-244-821-38

Alignment Scores:  
 Pred. No.: 1.59e-94 Length: 771  
 Score: 994.00 Matches: 193  
 Percent Similarity: 82.95% Conservative: 21  
 Best Local Similarity: 74.81% Mismatches: 26  
 Query Match: 40.72% Indels: 18  
 Gaps: 5

US-09-596-774-6 (1-461) x US-10-244-821-38 (1-771)

19 GlnIleGlnLeuValGlnSerGlyProGlyLeuValProGlyGlnThrValIle 38  
 1 CAGGTGAAGTCGACGAGTCGTCGAGTGAAGACCGGGTGAGACCGTCAGATC 60  
 39 SerGlyValSerGlyTyrProPheThrPheSerGlyValSerGlyValIle 58  
 61 AGCTGCAAGGCTTCTGTTAATCTTCACTGTTGTTGTTGTTGTTGTTGTTGTT 120  
 59 ProGlyGlnGlyLeuValTyrPheMetGlyTyrIleAsnThrSerThrGlyIle 78  
 121 CCGGGCAAGGTTTAAAGTGGATGGCTGATTAACCAAAACGTTGAACCACTAT 180  
 79 AlaAspPhePheGlyTyrGlyPheAspPheSerLeuGlnThrSerAlaAsnThr 98  
 181 GTTGAAGAGTTTAAAGGTCGCTTCTTCTTGGAGACCTGCGCACCACTGCTAT 240  
 99 LeuGlnIleAsnAsnLeuValSerGlyLeuAspMetAlaThrTyrPheCysAlaArgTyr 118  
 241 TTGAGATCAACCACTTAAATGAGACCGGTTAATTTCTGCACTTGGAGAC 300  
 119 ValTyrHisGlyTyrVal-----ProTyrTyrGlyGlnGlyThrValThrVal 135  
 301 TTCTAT---GATTACGTGAAGCTATGATTAATGCGGCCCAAGGACCACTGCTC 357  
 136 Ser-----SerGlyGlyGlySerGly 143  
 358 TCCAAAGATCTGTCGCGGTGCTCGGGCGGTGCGGTGCGGTGCGGTGCGGTGCGGT 417  
 144 GlyGlyGlySerGlyGlyGly---SerAspIleGlnLeuThrGlnSerHisAsp 162  
 418 GGTGTGTGGTGGGGGGGGGGGGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 477  
 163 LeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSerGlnAspValTyr 182  
 478 ATGTCACCTTCAAGTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 537  
 183 AsnAlaValAlaTyrTyrGlnGlnLeuValProGlyGlnSerProValLeuLeuIleTyrSer 202  
 538 ACGAATGTTGCTGATCAACAGAAACCGGGTCAATCCCGAAAGCACTGATTTACG 597

203 AlaSerArgTyrThrGlyValProSerArgPheThrGlySerGlyProAsp 222  
 598 GCATCTACCTTACAGTGTGTCGGATGCTTCACTGCGGACAGTGTCTGGACGAT 657  
 223 PheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCysGln 242  
 658 TTCACGCTCACCAACAGAAATGATGATGATGATGATGATGATGATGATGATGATGAT 717  
 243 HisPheArgThrPro---PheThrPheGlySerGlyThrIleValIleLeu 259  
 718 TATTACACTATTCGTTATTCAGTGTGCGGTGCGGAGACCAAGTTGGAATGAAG 771

RESULT 8

US-09-887-853-1  
 Sequence 1, Application US/09887853  
 Patent No. US20020168375A1  
 GENERAL INFORMATION:  
 APPLICANT: Huston, James S.  
 Oppermann, Hermann  
 Houston, L. L.  
 Ring, David B.  
 TITLE OF INVENTION: Biosynthetic Binding Proteins For Imaging

NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESS: Teesta, Hurwitz & Thibault/Patent Department  
 STREET: Exchange Place, 53 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/887,853  
 FILING DATE: 21-Jun-2001  
 CLASSIFICATION: <unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/133,804  
 FILING DATE: <unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kelley, Robin D.  
 REGISTRATION NUMBER: 34,637  
 REFERENCE/DOCKET NUMBER: 2054/22  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-248-7477  
 TELEFAX: 617-248-7100

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 909 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 3..752  
 OTHER INFORMATION: /product= "741F8 sFv' C-terminal Gly4-Cys"  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-887-853-1

Alignment Scores:  
 Pred. No.: 4.09e-93 Length: 909  
 Score: 981.50 Matches: 185  
 Percent Similarity: 85.89% Conservative: 22  
 Best Local Similarity: 76.76% Mismatches: 33  
 Query Match: 40.21% Indels: 1  
 Gaps: 1

DB: 9

US-09-596-774-6 (1-461) x US-09-887-853-1 (1-909)

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 38  
 Db 9 GAGATCCAAATGGTGGACGTGACCTGACCTGAAGAGCTGGAGAGACGTCAGATC 68

QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 58  
 Db 69 TCCTGCAAGGCTTCTGGGTATACCTTCAACAACTATGAAATGAACCTGGGTGAAGCGGCT 128

QY 59 ProGlyGlnGlyLeuLysTyrPheGlyTyrPheLeuAsnThrSerThrGlyGlnSerThrPhe 78  
 Db 129 CCAGGAAAGGGTTTAAAGTGGATGGCTGGATAAACCAACCACTGAGAGCCAACTAT 188

QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 98  
 Db 189 GCTGAAGAAGTTCAAGGAGCGGTTTGGCTTCTCTTGGAAACCTTCGACGCTGCTCAT 248

QY 99 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaIleTyrGlu 118  
 Db 249 TTGCAGATCAACAACCTCAAAATGAGGACAGCGCTACATATTTCTGTGAAGGCAATT 308

QY 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 138  
 Db 309 ATTACCTACGGCGGGTTTGCTAACTGGGGCCAAAGGACTCTGGTCACTGCTCTGCA 365

QY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158  
 Db 366 TCGAGCTCCGCGGATCTTCATCTGACGCTTCAGCTGAGCATATCTCATGTAGCCAG 425

QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178  
 Db 426 TCTCCTAAATTCATGCCACGTGAGTGGAGACAGGGTCAGCATCTCCGCAAGGCCAGT 485

QY 179 GlnAspValTyrAsnAlaValAlaTyrPyrGlnGlnLysProGlyGlnSerProLysLeu 198  
 Db 486 CAGATGTAGTACGTGCTTACCTCGTATCAACAAACCAAGGCAATCTCTCAAACTA 545

QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
 Db 546 CTGATTTACTGACATCCACCGGACACCTGGAGTCCCTGATCGCTTACAGGCAAGTGA 605

QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238  
 Db 606 TCTGGGACAGATTAATCTCACTCACTCAGAGTGTGAGGCTGGAAGCTGGCACTTCAT 665

QY 239 PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyTyrThrLysLeuGluIle 258  
 Db 666 TACTGTACGCAACATTATAGATGCGGTACACGTTCCGAGGGGGACCAAGCTGGAGATA 725

QY 259 Lys 259  
 Db 726 AAA 728

RESULT 9  
 US-10-683-547-1  
 ; Sequence 1, Application US/10683547  
 ; Publication No. US20050058638A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Huston, J.  
 ; APPLICANT: Houston, L.L.  
 ; APPLICANT: Ring, D.  
 ; APPLICANT: Oppermann, H.  
 ; TITLE OF INVENTION: BIOSYNTHETIC BINDING PROTEINS FOR IMMUNO-TARGETING  
 ; FILE REFERENCE: CIBT-P01-110  
 ; CURRENT APPLICATION NUMBER: US/10/683,547  
 ; PRIOR APPLICATION NUMBER: US/09/558,741  
 ; PRIOR FILING DATE: 2000-04-26  
 ; PRIOR APPLICATION NUMBER: 07/831,967  
 ; PRIOR FILING DATE: 1992-02-06  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn version 3.1

SEQ ID NO. 1  
 ; LENGTH: 909  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: 741F8 sfv'  
 ; NAME/KEY: CDS  
 ; LOCATION: (3)..(752)  
 ; OTHER INFORMATION:  
 US-10-683-547-1

Alignment Scores:  
 Pred. No.: 7,47e-92  
 Score: 969.50  
 Percent Similarity: 85.06%  
 Best Local Similarity: 75.93%  
 Query Match: 39.72%  
 DB: 19 Gaps: 1

US-09-596-774-6 (1-461) x US-10-683-547-1 (1-909)

QY 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 38  
 Db 9 GAGATCCAAATGGTGGACGTGACCTGACCTGAAGAGCTGGAGAGACGTCAGATC 68

QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 58  
 Db 69 TCCTGCAAGGCTTCTGGGTATACCTTCAACAACTATGAAATGAACCTGGGTGAAGCGGCT 128

QY 59 ProGlyGlnGlyLeuLysTyrPheGlyTyrPheLeuAsnThrSerThrGlyGlnSerThrPhe 78  
 Db 129 CCAGGAAAGGGTTTAAAGTGGATGGCTGGATAAACCAACCACTGAGAGCCAACTAT 188

QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 98  
 Db 189 GCTGAAGAAGTTCAAGGAGCGGTTTGGCTTCTCTTGGAAACCTTCGACGCTGCTCAT 248

QY 99 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaIleTyrGlu 118  
 Db 249 TTGCAGATCAACAACCTCAAAATGAGGACAGCGCTACATATTTCTGTGAAGGCAATT 308

QY 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 138  
 Db 309 ATTACCTACGGCGGGTTTGCTAACTGGGGCCAAAGGACTCTGGTCACTGCTCTGCA 365

QY 139 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158  
 Db 366 TCGAGCTCCGCGGATCTTCATCTGACGCTTCAGCTGAGCATATCTCATGACCCAG 425

QY 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178  
 Db 426 TCTCCTAAATTCATGCCACGTGAGTGGAGACAGGGTCAGCATCTCTCAAGGCGAGT 485

QY 179 GlnAspValTyrAsnAlaValAlaTyrPyrGlnGlnLysProGlyGlnSerProLysLeu 198  
 Db 486 CAGATGTAGTACGTGCTTACCTCGTATCAACAAACCAAGGCAATCTCTCAAACTA 545

QY 199 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 218  
 Db 546 CTGATTTACTGACATCCACCGGACACCTGGAGTCCCTGATCGCTTACAGGCAATGGA 605

QY 219 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 238  
 Db 606 TCTGGGACAGATTAATCTCACTCACTCAGAGTGTGAGGCTGGAAGCTGGCACTTCAT 665

QY 239 PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyTyrThrLysLeuGluIle 258  
 Db 666 TACTGTACGCAACATTATAGATGCGGTACACGTTCCGAGGGGGGACCAAGCTGGAGATA 725

QY 259 Lys 259  
 Db 726 AAA 728



RESULT 10  
US-08-812-393A-1  
Sequence 1, Application US/08812393A  
Publication No. US20010007152A1  
GENERAL INFORMATION:  
APPLICANT: SHERMAN, Linda A.  
APPLICANT: LUSTGARTEN, Joseph  
TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING  
TITLE OF INVENTION: T CELL RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR  
TITLE OF INVENTION: ANTIGENS  
NUMBER OF SEQUENCES: 64  
CORRESPONDENCE ADDRESSES:  
ADDRESSEE: MORRISON & FOERSTER  
STREET: 2000 Pennsylvania Avenue, NW, suite 5500  
CITY: Washington  
STATE: DC  
COUNTRY: USA  
ZIP: 20006-1888  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Diskette  
COMPUTER: IBM Compatible  
OPERATING SYSTEM: DOS  
SOFTWARE: FASTSEQ for Windows Version 2.0  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/812,393A  
FILING DATE: 05-MAR-1997  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER:  
FILING DATE:  
ATTORNEY/AGENT INFORMATION:  
NAME: Mureshige, Kate H  
REGISTRATION NUMBER: 29,959  
REFERENCE/DOCKET NUMBER: 31333-20001.00  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 202-887-1500  
TELEFAX: 202-822-0168  
TELEX:  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1350 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: double  
TOPOLOGY: linear  
FEATURE:  
NAME/KEY: Coding Sequence  
LOCATION: 1...1332  
OTHER INFORMATION:  
US-08-812-393A-1  
Alignment Scores:  
Pred. No.: 1,82e-87 Length: 1350  
Score: 930.00 Matches: 217  
Percent Similarity: 59.33% Conservative: 50  
Best Local Similarity: 48.22% Mismatches: 145  
Query Match: 38.10% Indels: 38  
DB: 8 Gaps: 14  
US-09-596-774-6 (1-461) x US-08-812-393A-1 (1-1350)  
QY 21 GlnueuValGlnserGlyProGlu-----LeuylsYsProGlyGluThrValylsle 38  
DB 70 CAGCAAGGAGCAGACAGTCCCGCATCTTGTCTGCAAGAGGAGAGAGCAGAGCTC 129  
QY 39 SerCyAlaSerGlyTyreProPheThraenTyrglyMetAsnTyrValysGlnAla 58  
DB 130 CAGGTACTTTTC-----ATCTTACAAACAG---GTGCATGCTTTTACCAAGCT 180  
QY 59 ProGlyGlnGlyLeuylsTyrPheTyrPheTyrPheTyrPheTyrPheTyrPhe 78  
DB 181 CCGGGGAGAGACCTGTCAGCCTGTTGTACATCCTCTCGGACAAACAGAGT----- 234

QY 79 AlaAspAspPheylsGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98  
DB 235 -----GGGAGACTGATCATCAACAGAGTCAATTAAGAGAGTGCAGCTCT 279  
QY 99 LeuGlnIleAsnAsnLeuylsSerGluAspMetAlaThrTyrPheCyAlaArgTyrGlu 118  
DB 280 TTGCATTTCTCTCTCCAGATCAACAGACTCAGGCACTTATCTCTGCTCAAAATCT 339  
QY 119 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 138  
DB 340 GGAGAGCAATTCGAAGCTTAACCTTCGGGAAAGCACTTAACCTCTGTTAAATCAGT 399  
QY 139 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158  
DB 400 GCGGAGAGGTCTGGCGGGGGTGGATCCGGGGTGGAGGCTCAGAGGCTGAGTCAACCCAA 459  
QY 159 SerHisYsPheLeuSerThrSerValGlyAspArgValSerIleThrCyAlaAlaSer 178  
DB 460 AGCCCAAGAAACAAAGTGGCAGTAAACAGAGAAAGTGACATTGAGCTTAATCAGACT 519  
QY 179 GlnAspValTyrAsnAlaValAlaTyrPheGlnGlnYsProGlyGlnSerProYsAsp 198  
DB 520 AATAAC--CAACAACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 576  
QY 199 LeuIleTyrSer-----AlaSerSerArgTyrThrGlyValProSerArgPheThr 215  
DB 577 ATCATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTATTTAT 636  
QY 216 GlySer---GlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 234  
DB 637 GCCTCCAGACCAACCAAGAGAACTTCCCTCATTTCTGAGATTGGTACCCTCTCAG 696  
QY 235 LeuAlaValTyrPheCys-----GlnGlnHisPheArgThrProPheThrPheGly 251  
DB 697 ACATCAGTACTTCTGTGCGCAGCGGTAGACAGGACCAACAAAGATTTATTTTCGT 756  
QY 252 SerGlyThrIleLeuGluIleLeuAlaLeuGluIleSerAsnSerValMetTyrPheSer 271  
DB 757 CATGGAACCAAGCTGTCTGCTCAGT-----AGTAACTCATCATGATTCATTCAGC 807  
QY 272 SerValValProValLeuGlnLeuValAsnSerThrThrIleValProValLeuArgThr 291  
DB 808 CACTTCGTCGCGCTCTCTCCAGCAGAGAGCCCAACAGAGC---CCAGCCCGCGAGCA 864  
QY 292 ProSerProValHisProThrGlyThrSerGlnProGlnArgProGluAspCysArgPro 311  
DB 865 CCAACACCGCGC---CCACCATCGCGTGCAGCGCCCTG-----TCCCTGCGCCCA 912  
QY 312 ArgGlySerValIleGlyThrGlyLeuAspPheLeuGluAspProYsLeuCyAlaTyrLeu 331  
DB 913 TCTAGTTCT-----AGAGATCCCAAACTGTGTAAGTCTG 945  
QY 332 LeuAspGlyIleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaYs 351  
DB 946 CTGATGGAATCTCTTCATCTATGAGTGTCACTTCATCTGCTGCTGCTGCTGCTGCTG 1005  
QY 352 PheSerArgSerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlu 371  
DB 1006 TTCAGCAGAGCAGAGAGCCCGCCGCTAACAGAGAGGCGCAACCACTCTATTAACAG 1065  
QY 372 LeuAsnLeuGlyArgArgGluGluTyrAspValLeuGluYsYsArgAlaArgAspPro 391  
DB 1066 CTCAATCTAGACCAAGAGAGAGTACATTTTGGACAAAGACGTCGCGGAGACCTT 1125  
QY 392 GluMetGlyGlyGlyGlnGlnArgArgArgAsnProGlnGlnGlyValTyrAsnAlaLeu 411  
DB 1126 GAGATGGGGGAGAAAG---CCAGAGAGAGAAACCTCAGAAAGCCTGTACAAATGAAG 1182  
QY 412 GlnYsAspYsMetAlaGluAlaTyrSerGluIleGlyThrIleYsGlyGluArgArgArg 431  
DB 1183 CAGAAAGTAAAGATGGCGAGGCTTACAGTGAATGGAGTGAAGGAGGAGCGCGGAGG 1242  
QY 432 GlyYsGlyHisAspGlyLeuTyrGlnGlyLeuSerThrAlaThrIleYsAspThrTyrAsp 451

Db 1243 GGCAGAGGGCAGCATGGCTTTTACAGAGGCTCTCAGTACACCCACCAAGACCTTACGAC 1302

Qy 452 AlaleuHismetGlnThrLeuAlaProArg 461  
Db 1303 GCCCTTCACATGACGGCCCTGCCCCCTCGC 1332

## RESULT 11

US-09-774-681-1  
; Sequence 1, Application US/09774681  
; Publication No. US20030208780A1  
; GENERAL INFORMATION:  
; APPLICANT: Sunol Molecular Corporation  
; APPLICANT: Sherman, Linda  
; APPLICANT: Lustgarten, Joseph  
; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL  
; FILE REFERENCE: 31333-20001.01  
; CURRENT APPLICATION NUMBER: US/09/774,681  
; PRIOR FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: US 08/812,393  
; PRIOR FILING DATE: 1997-03-05  
; PRIOR APPLICATION NUMBER: US 60/012,845  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1350  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide derivative of effective T cell  
; NAME/KEY: CDS  
; LOCATION: (1)...(1350)  
US-09-774-681-1

Alignment Scores:  
Pred. No.: 2,956-87 Length: 1350  
Score: 928.00 Matches: 217  
Percent Similarity: 59.33% Conservative: 50  
Best Local Similarity: 48.22% Mismatches: 145  
Query Match: 38.02% Gaps: 38  
DB: 10 Indels: 14

US-09-596-774-6 (1-461) x US-09-774-681-1 (1-1350)

Qy 21 GlnLeuValGlnSerGlyProGlu-----LeuYsLysProGlyGluThrValLysIle 38  
Db 70 CAGCAAGTGCAGCAGATCCCGCATCTTGGTTCGACGAGGGGAGAAACGACAGCTC 129  
Qy 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 58  
Db 130 CAGGTGATCTTTCC-----ATCTTACAAACAG--GTGAGTGGTTTACCAACGT 180  
Qy 59 ProGlyGlnGlyLeuLysTyrPheMetGlyTyrPheLeuAsnThrSerThrGlyLysThrPhe 78  
Db 181 CTTGGGGGAAAGCTCTGACGCTGTTTACATCTTCTGGCAAGGAGAGT-----234  
Qy 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 98  
Db 235 -----GGAGACTGCATCCACAAAGTCACTTAAAGAACTCGCAGCTC 279  
Qy 99 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 118  
Db 280 TTGCACATTTCTCTCTCCAGATCACAGACTCAGGACCTATCTCTGCTCAATCTC 339  
Qy 119 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 138  
Db 340 GGAAGAAACATGACAAAGCTACCTTCGGGAAAGGACCTAACTCTCTGTTAAATCAGGT 399  
Qy 139 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 158

Db 400 GCGGAGGCTGTGCGGGGGTGATCCGGGGTGAGGCTCAGAGCTGCAGTCAACCCAA 459  
Qy 159 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 178  
Db 460 AGCCCAAGAAACAAAGGTGGCGATACACAGAGAAAGGTGACATTTGATTAATCAGACT 519  
Qy 179 GlnAspValTyrAsnAlaValAlaTyrPyrGlnGlnLysProGlyGlnSerProLysLeu 198  
Db 520 AATAAC---CACACAACATGTAAGTGTATCGGAGAGACAGGGGCAATGGCGAGGCTG 576  
Qy 199 LeuIleTyrSer-----AlaSerSerArgTyrThrGlyValProSerArgPheThr 215  
Db 577 ATCCATTATTCATATGCTGTCTGCGACGACAGAAAGAGATATCCCTGATGATCAAG 636  
Qy 216 GlySer---GlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 234  
Db 637 GCGTCCAGACCAAGCCCAAGAAAGTCTCTCCCTCATTTGTGAGTTGGGTACCCCTTCAG 696  
Qy 235 LeuAlaValTyrPheCys-----GlnGlnHisPheArgThrProPheThrPheGly 251  
Db 697 ACATCAGTGTACTTCTGTGCGACGGGTGAGACAGGACCAACGAAGATTATTTTCGGT 756  
Qy 252 SerGlyThrLysLeuGlnIleLysAlaLeuGlnIleSerAsnSerValMetTyrPheSer 271  
Db 757 CATGAAACCAAGCTGTCTGTCTGACT-----AGTAACTCCATCATATGACTTCAGC 807  
Qy 272 SerValValProValLeuGlnLysValAsnSerThrThrLysProValLeuArgThr 291  
Db 808 CACTTCGTGCGGTCTCTCTGCGACGAGGACCAAGCCACACAGACG---CCAGCGCGGAGCA 864  
Qy 292 ProSerProValHisProThrGlyThrSerGlnProGlnArgProGlyLysArgPro 311  
Db 865 CCAACACCCGCG---CCACACATCGCGTGCAGCCCTCG-----TCCCTGGGCCCA 912  
Qy 312 ArgGlySerValLysGlyThrGlyLeuAspPheLeuGluAspProLysLeuCysTyrLeu 331  
Db 913 TCTAGTTCT-----AGAGATCCCAACTCTGCTAAGCTG 945  
Qy 332 LeuAspGlyIleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLys 351  
Db 946 CTGAGAGGATCTCTTCATCTATGATGTCATTCATCTGCTGCTTCTGAGATGAAG 1005  
Qy 352 PheSerArgSerAlaGlnThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlu 371  
Db 1006 TTCAGAGAGGCGCAGACCCCGCGTACACAGCGGCGCAAGACCACTCATTAACGAG 1065  
Qy 372 LeuAsnLeuGlyValArgGlyGlnIleTyrAspValLeuGlnLysArgAlaArgAspPro 391  
Db 1066 CTCATCTTAGACGAAAGAGAGATGATGTTTGGACAAAGACGTGGCGGGACCT 1125  
Qy 392 GlnMetGlyGlyLysGlnGlnLysArgArgAsnProGlnGlnGlyValTyrAsnAlaLeu 411  
Db 1126 GAGATGGGGGGAAG---CCGAGAAAGAAAGACCTTCGAGAGGCTGTACAAAGAACTG 1182  
Qy 412 GlnLysAspLysMetAlaGlnAlaTyrSerGlnIleGlyThrLysGlyGlnLysArgArg 431  
Db 1183 CAGAAAGATTAAGATGGCGAGGCTTACAGATGAGATGGAGAAAGGAGCGCCGAGAG 1242  
Qy 432 GlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAsp 451  
Db 1243 GCGAAGGGGACAGATGGGCTTTACAGGCTCTCAGTACAGCAACAAAGACACCTACGAC 1302  
Qy 452 AlaleuHismetGlnThrLeuAlaProArg 461  
Db 1303 GCCCTTCACATGACGGCCCTGCCCCCTCGC 1332

## RESULT 12

US-10-239-656-78  
; Sequence 78, Application US/10239656  
; Publication No. US20040038339A1  
; GENERAL INFORMATION:  
; APPLICANT: KUPFER, PETER  
; APPLICANT: RIETHMULLER, GERT

APPLICANT: LUTTERBUSE, RALF  
 APPLICANT: BORSCHERT, KATRIN  
 APPLICANT: KISCHEL, ROMAN  
 APPLICANT: MAYER, MONIKA  
 APPLICANT: HOMMEISTER, ROBERT  
 TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE  
 TITLE OF INVENTION: TO AN EPITOPE OF THE NK2D RECEPTOR COMPLEX  
 FILE REFERENCE: 029976/0106  
 CURRENT APPLICATION NUMBER: US/10/239, 656  
 PRIOR APPLICATION NUMBER: PCT/EP01/03414  
 PRIOR FILING DATE: 2001-03-26  
 PRIOR APPLICATION NUMBER: EP 00106467.4  
 PRIOR FILING DATE: 2000-03-24  
 NUMBER OF SEQ ID NOS: 92  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO 78  
 LENGTH: 1515  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic 3810xPS-  
 US-10-239-656-78  
 OTHER INFORMATION: 23 bp specific single chain fv

Alignment Scores:  
 Pred. No.: 4.38e-86 Length: 1515  
 Score: 917.50 Matches: 190  
 Percent Similarity: 56.99% Conservative: 30  
 Best Local Similarity: 49.22% Mismatches: 38  
 Query Match: 37.59% Gaps: 129  
 DB: 17 Indels: 5

US-09-596-774-6 (1-461) x US-10-239-656-78 (1-1515).

2 AlaTTPValTTPThrLeuLeuPheLeuMetAlaAlaAlaLys-----ValProLys 18  
 323 AGCTGGAAATCAAGGTGGTGGTCTCTGGCGGGCGGCGCTCGGTGGTGGTGTCT 381  
 19 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGlnThyValLysIle 38  
 382 CAGGTGCAACTGCGACGACTGACGGCTGAGCTGAAGAAAGCTCGAGAGACAGTCAAGATC 441  
 39 SerCysValSerAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTTPValLysGlnAla 58  
 442 TCTGCGAAGGCTTCTGGGTATACCTTCACAAACTATGGAATGAATGAGGTAAAGCAGGCT 501  
 59 ProGlyGlnGlyLeuLysTTPMetGlyTTPIleAsnThrSerThrGlyLysSerThrPhe 78  
 502 CCAGGAAAGGCTTTCACAGTGGATGGCTGGATTAACACTACACTGAGAGACCAACATAT 561  
 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 98  
 562 GGTATGACTTCAACGGGAGCGGTGGCTCTTCTTGGAAACCTCGCCAGACATGCTAT 621  
 99 LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTyrPheCysAlaArg----- 116  
 622 TTGCGATCAACAACTCAAAAATGAGGACACGGCTACATATTTCTGCAAGATTCAAC 681  
 116 ----- 116  
 682 TCCCTGACTACTGGGCGCAAGGACACAGGTCAACCGTCTCTCCGAGAGTGGTGAATCC 741  
 116 ----- 116  
 742 GAGGTGACGTGCTGAGTCTGAGAGTGGCTGTGTGACAGCTGAGAGATCCCTGAACCTC 801  
 116 ----- 116  
 802 TCTGTGACGCTCAGAGTTGATTTTGTAGATACCTGAGTGAATGGGTCCGCGAGGCT 861  
 117 ----- 117

Db 862 CCAGGAAAGGCTAGAAATGATGGAGAAATTAATCCAGATAGCATAGATAAATAT 921  
 117 ----- 117  
 Db 922 AGCCATCTTAAGAGATAGATTATCATCTTCACAGACAAAGCCAAATAATGCTGTAC 981  
 118 ----- 121  
 Db 982 CTGCAATGACCAAGTAGAGTCTGAGGACACACCCCTTTATTTACTGTCGCAAGATTGGGG 1041  
 122 -----GlyTyrValProTyrTTPGlyGlnGlyThrThrValThrValSerSerGlyGly 139  
 1042 CAATGGGGTACTTGTGACTACTGGGGCCAAAGGACCAAGCTCACCGTCTCTCAGTGGT 1101  
 140 GlnGlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGlnSer 159  
 1102 GGTGTTCTGGCGGGCGGCTCGGTGGTGGTGTGAGCTGCTGATGATACACAGTCT 1161  
 160 HisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSerGln 179  
 1162 CCATCTCTCCCTGACTGTGACACAGGAGAGAGGCTCATATGAGTGCAGATCCAGTCA 1221  
 180 AspValTyrAsnAla-----ValAlaTTPTyrGlnGlnLysProGly 193  
 1222 AGTCTGTTAAACAGTGAATCAAAAGAACTACTGACTGTACCTGACACAGAAACACAGG 1281  
 194 GlnSerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThrGlnValProSerArg 213  
 1282 CAGCTCTTAACGTGTATCTACTGGGCATCTCAGAGAAATCTGGGGTCTCGATCC 1341  
 214 PheThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlu 233  
 1342 TTCACAGGACATGATCTGGAACAGATTTCACCTCCACCATCAGAGTGTCAAGCTGAA 1401  
 234 AspleuAlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGly 253  
 1402 GACCTGGAGATTATTAATCTGCAGAAATGATATATATCTCTACAGCTGTGGTGG 1461  
 254 ThrLysLeuGluLysLys 259  
 1462 ACCAAGCTTGATCAAA 1479

RESULT 13  
 US-09-766-543-9  
 Sequence 9, Application US/09766543  
 Patent No. US20020041865A1  
 GENERAL INFORMATION:  
 APPLICANT: Auestin, Richard  
 APPLICANT: Kwok, Cheuk S.  
 APPLICANT: Ring, David B.  
 TITLE OF INVENTION: METHODS FOR TREATING TUMORS  
 FILE REFERENCE: PP01679.002  
 CURRENT APPLICATION NUMBER: US/09/766,543  
 PRIOR FILING DATE: 2000-01-19  
 PRIOR APPLICATION NUMBER: 60/177,258  
 NUMBER OF SEQ ID NOS: 14  
 SOFTWARE: PatentIn Ver. 2.0  
 SEQ ID NO 9  
 LENGTH: 816  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: 520C9  
 OTHER INFORMATION: humanized single-chain antibody used in the  
 NAME/KEY: CDS  
 LOCATION: (7) (807)  
 US-09-766-543-9

Alignment Scores:  
 Pred. No.: 5.6e-86 Length: 816  
 Score: 913.00 Matches: 167

Percent Similarity:	83.61%	Conservative:	37
Best Local Similarity:	68.44%	Mismatches:	36
Query Match:	37.40%	Indels:	4
DB:	9	Gaps:	2

US-09-596-774-6 (1-461) X US-09-766-543-9 (1-816)

Qy 1.9 GlnIleGlnLeuValG

Conservative:	37
Mismatches:	36
Indels:	4
Gaps:	2

3-9 (1-816)

proglutinin vs pro

||| :: |||

Dp	3	GAGGATGCAACCTGGGAGGAGTCTGGGGCCCTGAGGTGANAAGAGCTGGGGCTCACTGAAAGGTC	1
Qy	39	SerCysValSerAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnThrValLeuGlnIle	58
Dp	133	TCCTGCAAGGCTTCTGGTTACACCTTTACCACTTACGATGAGTAACTGGGTGGCAAGGCGC	19
Qy	59	ProGlyGlnGlyLeuLysTrpMetGlyTrpIleAsnThrSerThrGlyLeuSerThrPhe	78
Dp	193	CCTGAGCAAGGCGCTTGATGGATGGATGATGATTAACCACTTACCTGACAGTCAACATAT	25
Qy	79	AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnIleThrSerAlaAsnThrValTyr	98
Dp	253	GCTATGATCTTCAAGGAAGAGTCAACCATGACCAAGACATCAACGACAGCAAGCCTTAC	31
Qy	99	LeuGlnIleAsnAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTrpIle	118
Dp	313	ATGAGACCTGAGAGCGCTGAGATCTGAGACAGACGCGGTATATTACTGTGCGAGA-----	36
Qy	119	ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly	138
Dp	367	---CGATTTTGAGTTTGCT---TACTGGGGCGAGGAAACCTGTGACCGTCTCTCAGGT	42
Qy	139	GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln	158
Dp	421	GGCGGTGGCTCGGGCGGTGGGTGGGTGGGTGGCGGCGGATCTGACATCAAGATGACCGAG	48
Qy	159	SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaAlaSer	178
Dp	481	TTCTCCATCCCTCGTCTGTCATCTGAGAGACAAGTCAACATCATCTTGGCGGGCAAGT	54
Qy	179	GlnAspValTyrAsnAlaValAlaTyrGlnIleLysProGlyGlnSerProLysLeu	198
Dp	541	CAGGACATGTGGTAATAGCTTAACCTGTATCCAGCAAAACAGGAAAAACCCCTTAACCTC	60
Qy	199	LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly	218
Dp	601	CTATCTACCCCACTCCAGTTAGATTCTGGGGGTCCCATCAAGTTCACTGAGAAAGTGA	66
Qy	219	SerGlyProAspPheThrPheThrIleSerSerValGlnIleGluAspLeuAlaValTyr	238
Dp	661	TCTGGAGCAAGATTTTACCTTACCAATCAAGCAATCTGCGCTGGAAGATATTGGCAACATAT	72
Qy	239	PheCysGlnGlnHisIleAsnArgThrProPheThrPheGlySerGlyTyrTrpLeuGlnIle	258
Dp	721	TACTGTCTCAATATGCTATTTTTCGTACACGTTTGGCGCAAGGAGACACATCTGAGATT	78
Qy	259	LysAlaLeuGlu	262
Dp	781	AAAGATCCGAA	792

QY 119 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValSerSergly 138  
 DB 298 GGTACACGCTGTAAGTCTGATGTCGGGGCCAGGAGCACCGTCTCCCTCAGGT 357  
 QY 139 GlyGlyGlySerglyGlyGlyGlySerglyGlyGlySeraPileGlnLeuThrGln 158  
 DB 358 GCGCGTGGATCTGGGAGGAGGCGGCGAGGAGGATGATCATCAGATGACTCAG 417  
 QY 159 SerHisLysPheLeuSerThrSeraValGlyAspArgValSerLileThCysLysAlaSer 178  
 DB 418 TCTCATCTTCCCTGCTCTGATCTCTAGAGACAGAGTCACTATCACTTGCAGGCGAGT 477  
 QY 179 GlnAspValTyrAsnAlaValAlaTyrPyrGlnGlnLysProGlyGlnSerProLysLeu 198  
 DB 478 CAGGACATTAATAGCTATTAAAGCTGGTTCCAGCAAGAAACCGGAAAGCTCTTAAGCC 537  
 QY 199 LeuIleTyrSeraLysSerSeraTyrGlyThrGlyValProSeraArgPheThrGlySergly 218  
 DB 538 CTGATCTATCGTGCAACACAGATGAGATGCGGCGTCCCATCAAGGTTCACTGCGCAGTGA 597  
 QY 219 SerGlyProAspPheThrPheThrLleSeraSerValGlnAlaGlyAspLeuAlaValTyr 238  
 DB 598 TCTGGACAGATTAATCTCTCACTCCATCAGCAGCTGCAATATGAAGATTGGAAATTTAT 657  
 QY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerglyThrLysLeuGluLile 258  
 DB 658 TATGTCTACACGATGATGATGATCTCCGTGAGCGTTCCGTGAGAGCACCAAGCTTGAGATG 717  
 QY 259 Lys 259  
 DB 718 AAA 720

RESULT 15

US-10-717-243-90  
 : Sequence 90, Application US/10717243  
 : Publication No. US20050054835A1  
 : GENERAL INFORMATION:

APPLICANT: Better, Marc D.  
 Carroll, Stephen F.  
 Studnicka, Gary M.

TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
 Proteins

NUMBER OF SEQUENCES: 169

CORRESPONDENCE ADDRESS:

ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor

CITY: Chicago

STATE: Illinois

COUNTRY: USA

ZIP: 60661

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patent Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/717,243

FILING DATE: 18-NOV-2003

CLASSIFICATION: 530

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/839,765

FILING DATE: 15-APR-1997

APPLICATION NUMBER: US 08/425,336

FILING DATE: 18-APR-1995

APPLICATION NUMBER: US 08/064,691

FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/988,430

FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787,567

ATTORNEY/AGENT INFORMATION:  
 NAME: McNicholas, Janet M.  
 REGISTRATION NUMBER: 32,918  
 REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 312/707-8889  
 TELEFAX: 312/707-9155  
 INFORMATION FOR SEQ ID NO: 90:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 723 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: DNA  
 SEQUENCE DESCRIPTION: SEQ ID NO: 90:  
 US-10-717-243-90

US-09-596-774-6 (1-461) x US-10-717-243-90 (1-723)

Alignment Scores:  
 Pred. No.: 4,73e-85 Length: 723  
 Score: 903.50 Matches: 167  
 Percent Similarity: 82.57% Conservative: 32  
 Best Local Similarity: 69.29% Mismatches: 41  
 Query Match: 37.01% Indels: 1  
 Gaps: 1

QY 19 GlnIleGlnLeuValGlnSerglyProGlyLeuLysPyrGlyGluThrValLysIle 38  
 DB 1 GAGATCCAGTGGTGGAGTCTGAGAGAGGCGCTGTGAAGCTTGAGAGGCTCGTCAGAAATC 60  
 QY 39 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValGlnAla 58  
 DB 61 TCTGCGCAGCTTGTGGATATACCTTCAACAATATGAAATGAGTGGTCCGCGAGCT 120  
 QY 59 ProGlyGlnGlyLeuLysTrpMetGlyTyrPheAsnThrSeraThrGlyGluSerThrPhe 78  
 DB 121 CCAGGAAAGGTTTAAAGTGTATGGCTGATTAACCCACCTGAGAGCCCAACATAT 180  
 QY 79 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSeraLysAsnThrAlaTyr 98  
 DB 181 GCTATTCCTTCAAGGAGACGCTTACCTTCTTTGAGACGATCTTAAGAACACCTGCCAT 240  
 QY 99 LeuGlnIleAsnLeuLysSeraGlyAspMetAlaThrTyrPheCysAlaArgTrpGlu 118  
 DB 241 TTACAGATCAACACGCTCAGACCGAGAGACCGCTGTATTTCTGTACAAGA---CGG 297  
 QY 119 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSergly 138  
 DB 298 GGTACACGCTGTAAGTCTGATGTCGGGGCCAGGAGCACCGTCTCCCTCAGGT 357  
 QY 139 GlyGlyGlySerglyGlyGlyGlySerglyGlyGlySeraPileGlnLeuThrGln 158  
 DB 358 GCGCGTGGATCTGGGAGGAGGCGGCGAGGAGGATGATCATCAGATGACTCAG 417  
 QY 159 SerHisLysPheLeuSerThrSeraValGlyAspArgValSerLileThCysLysAlaSer 178  
 DB 418 TCTCATCTTCCCTGCTCTGATCTCTAGAGACAGAGTCACTATCACTTGCAGGCGAGT 477  
 QY 179 GlnAspValTyrAsnAlaValAlaTyrPyrGlnGlnLysProGlyGlnSerProLysLeu 198  
 DB 478 CAGGACATTAATAGCTATTAAAGCTGGTTCCAGCAAGAAACCGGAAAGCTCTTAAGCC 537  
 QY 199 LeuIleTyrSeraLysSerSeraTyrGlyThrGlyValProSeraArgPheThrGlySergly 218  
 DB 538 CTGATCTATCGTGCAACACAGATGAGATGCGGCGTCCCATCAAGGTTCACTGCGCAGTGA 597  
 QY 219 SerGlyProAspPheThrPheThrLleSeraSerValGlnAlaGlyAspLeuAlaValTyr 238  
 DB 598 TCTGGACAGATTAATCTCTCACTCCATCAGCAGCTGCAATATGAAGATTGGAAATTTAT 657  
 QY 239 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerglyThrLysLeuGluLile 258

Db	658	TATTGTCAACGATATGATGAGTCTCCGTCGACGTTCCGTGGAGGACCACTTGAGATG	717
Qy	259	Lys	259
Db	718	AAA	720

Search completed: April 20, 2005, 17:07:32  
Job time : 812.902 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:53:52 ; Search time 17.8485 Seconds  
(without alignments)  
2485.142 Million cell updates/sec

Title: US-09-596-774-6  
Perfect score: 2441  
Sequence: 1 MAMWTLLFLMAAAKVPKQI.....ISTATKDTYDALHMOTLAPR 461

Scoring table: BLOSUM62  
Gapop 10.0 , Gapept 0.5

Searched: 283416 seqs, 96216763 residues  
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR\_79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	713	29.2	164	2 A40104	T-cell receptor CD
2	694	28.4	268	2 A56446	Ig heavy chain V r
3	665.5	27.3	249	2 S41374	single chain Fv an
4	611.5	25.1	233	2 JCS322	p53 specific singl
5	605	24.8	177	2 S54817	T-cell receptor lo
6	605	24.8	206	2 A35900	T-cell receptor CD
7	604.5	24.8	163	2 A31768	T-cell receptor ze
8	589.5	24.1	188	2 A45089	T-cell receptor CD
9	588.5	24.1	136	2 S35759	BHD9D10 protein -
10	585	24.0	146	4 S33905	Ig heavy chain pre
11	555	22.7	139	2 PH1225	Ig heavy chain pre
12	539	22.1	166	2 JG4654	T-cell receptor ze
13	539	22.1	166	2 T46424	T-cell surface gly
14	537.5	22.0	120	2 B42848	L6 mab heavy chain
15	536	22.0	124	2 PH1404	Ig heavy chain V r
16	522.5	21.4	118	2 S37204	Ig heavy chain V r
17	519	21.3	119	2 A53285	Ig heavy chain V a
18	516	21.1	119	2 H45722	anti-glycoprotein
19	510	20.9	117	2 S32190	Ig heavy chain V r
20	499.5	20.5	115	2 S19965	Ig heavy chain V r
21	495	20.3	107	2 S32192	Ig kappa chain V r
22	490	20.1	107	2 S32191	Ig kappa chain V r
23	490	20.1	113	2 B36259	Ig heavy chain V r
24	488.5	20.0	118	2 A32530	Ig heavy chain V r
25	486.5	19.9	115	2 S19968	Ig heavy chain V r
26	486	19.9	117	2 S32187	Ig heavy chain V r
27	486	19.9	118	2 S19967	Ig heavy chain V r
28	486	19.9	149	1 KWSM11	Ig kappa chain pre
29	484.5	19.8	109	2 S26325	Ig heavy chain V r

30	484.5	19.8	120	2 S19963	Ig heavy chain V r
31	484	19.8	119	2 B32530	Ig heavy chain V r
32	481.5	19.7	134	2 S21916	Ig heavy chain V r
33	480	19.7	119	2 P00265	Ig kappa chain V r
34	477.5	19.6	114	2 D32967	Ig heavy chain V r
35	472.5	19.4	114	2 C32967	Ig heavy chain V r
36	470.5	19.3	142	2 S19245	Ig heavy chain pre
37	469.5	19.2	105	2 S24765	Ig heavy chain V r
38	469	19.2	152	2 S30751	Ig kappa chain pre
39	460	18.8	114	2 P10256	Ig heavy chain V r
40	458.5	18.8	105	2 S24766	Ig heavy chain V r
41	457.5	18.7	105	2 S24764	Ig heavy chain V r
42	453.5	18.6	108	2 P10083	Ig kappa chain V r
43	452	18.5	117	2 S42466	Ig kappa chain V r
44	451.5	18.5	105	2 S24763	Ig heavy chain V r
45	451	18.5	131	2 S26792	Ig heavy chain V r

## ALIGNMENTS

RESULT 1  
A40104  
T-cell receptor CD3 zeta chain - mouse  
C:Species: Mus musculus (house mouse)  
C>Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 09-Jul-2004  
C:Accession: A40104; I55293  
R:Weisman, A.M.; Baniyash, M.; Hou, D.; Samuelson, L.E.; Burgess, W.H.; Klausner, R.D.  
Science 239, 1018-1021, 1988  
A:Title: Molecular cloning of the zeta chain of the T cell antigen receptor.  
A:Reference number: A40104; MUID:88145643; PMID:3278377  
A:Accession: A40104  
A>Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-164 <WEB>  
A:Cross-references: UNIPROT:P24161; GB:M19729; NID:G201133; PIDN:AAA40171.1; PID:G201132  
R:Baniyash, M.; Hsu, V.W.; Seldin, M.F.; Klausner, R.D.  
J. Biol. Chem. 264, 13252-13257, 1989  
A:Title: The isolation and characterization of the murine T cell antigen receptor zeta c  
A:Reference number: I55293; MUID:88327299; PMID:2787796  
A:Accession: I55293  
A>Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-164 <RES>  
A:Cross-references: GB:J04967; NID:G556326; PIDN:AAA50301.1; PID:G556327  
C:Gene: Tcrz  
C:Keywords: phosphoprotein; T-cell receptor; transmembrane protein

Query Match 29.2% Score 713; DB 2; Length 164;  
Best local similarity 99.3%; Pred. No. 4e-42; Mismatches 1; Indels 0; Gaps 0;  
Matches 138; Conservative 0;

QY 323 LEDPKLCYLDGILFIYGVITITLALYLRAKFSRSSETANLQDPNQLYNELNGRREEDV 382  
DB LLDRLKLCYLDGILFIYGVITITLALYLRAKFSRSSETANLQDPNQLYNELNGRREEDV 85  
QY 383 LEKRRADPEWKGKQQRNRNPQEGVYNALQDKMAEAYSEITGGERRRGKHGGLYQGL 442  
DB LEKRRADPEWKGKQQRNRNPQEGVYNALQDKMAEAYSEITGGERRRGKHGGLYQGL 145  
QY 443 STATKDTYDALHMOTLAPR 461  
DB STATKDTYDALHMOTLAPR 164

RESULT 2  
A56446  
Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)  
C:Species: Mus musculus (house mouse)  
C>Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 16-Aug-1996  
C:Accession: A56446  
R:Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.

J. Biol. Chem 270, 7829-7835, 1995  
 A:Title: A high affinity digoxin-binding protein displayed on M13 is functionally identical to the cardiac glycoside binding protein of *Streptococcus*  
 A:Reference number: A56446; MUID:95229583; PMID:771873  
 A:Accession: A56446  
 A:Status: preliminary  
 A:Molecule type: mRNA  
 A:Residues: 1-268 <TAN>  
 A:Cross-references: GB:U20617  
 C:Keywords: heterodimer; Immunoglobulin

Query Match	28.4%	Score 694;	DB 2;	Length 268;
Best Local Similarity	55.2%;	Pred. No. 1.5e-40;		
Matches 133;	Conservative 43;	Mismatches 63;	Indels 2;	Gaps 2;

Qy		19	QIOLVSGPELKKRGSETTVKISCKASGYPTTNGMMWVQAQGGLKMMGMIINTSTGE	78
Dd		3	QVLQSSGGALVYRKGASVLTSCITTSGFNKLDTYHMVYQRREQGLEHIGRIAPANGITTK	62
Qy		79	ADDFKGRFDESLTETSANTAYIQINLKSEDMATYFCARMEVHYGVYPWGQTIVYS	138
Dd		63	DPFKOKARTIAADTSSNTAYILOSLITSBEDPVTAVYCASYLTR-YENVMGQTIVYS	121
Qy		139	GGSSGGSSGGSSDILQLTOSHKFLSTVGDRVSLTCASDYVNAAWTQQRGSPKL	198
Dd		122	GGSSGGSDSGGGSDIELTPALIMASISGEKVYTWSCASSSV-NFIYWQQKSDASP	180
Qy		199	LIIYSASSRYTGVPSPRTFSGSGPDFTFTTISSVOAEDLIWVYCCQHFRPFTFGS	258
Dd		181	WYYTYTHLPYGVARPSGSGSNYSLTISSMEGEDATATYCCQFTSPFTFGSKTLEI	240
Qy		259	K 259	
Dd		241	K 241	

### RESULT 3

single chain Fv antibody - mouse  
C;Species: Mus musculus (house mouse)  
C;Date: 06-Jan-1995 #sequence\_revision 06-Jan-1995 #text\_change 06-Jan-1995  
C;Accession: S41374  
R;Artseenko, O.; Weller, E. W.; Muentz, K.; Conrad, U.  
submitted to the EMBL Data Library, January 1994  
A;Description: Construction and functional characterization of a single chain Fv antibody  
A;Reference number: S41374  
A;Accession: S41374  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-249 <ART>  
A;Cross-references: EMBL:Z23480

Query Match	27.3%;	Score 665.5;	DB 2;	Length 249;
Best Local Similarity	52.0%;	Pred. No. 1.2e-38;		
Matches 128;	Conservative 47;	Mismatches 66;	Indels 5;	Gaps 1

[illegible]

Db 241 TKLELK 246

```

RESULT 4
JC5322
p53 specific single-chain antibody Pab421 - human
C:Species: Homo sapiens (man)
C:Date: 15-May-1997 #sequence_revision 15-May-1997 #text_change 18-Jul-1997
C:Accession: J05322

```

Query Match	25.1%	Score 61.5	DB 2	Length 233
Best Local Similarity	51.0%	Pred. No. 5.8e-35		
Matches 122, Conservative	39	Mismatches 67	Indels 11	Gaps 3

[illegible][illegible]

## RESULT 5

T-cell receptor Iota precursor - mouse  
C/Species: Mus musculus (house mouse)  
C/Date: 08-Jul-1995 #sequence\_revision 21-Jul-1995 #text\_change 05-Nov-1999  
C/Accession: S54817  
R/Ronchetti, G.; Ronchetti, S.; Baroli, A.; Testa, G.; d'Adamo, F.; Riccardi, C.; Migliorini, M.  
Submitted to the EMBL Data Library, January 1995  
A/Description: T cell receptor Iota: an alternatively spliced product of the T cell receptor gene  
A/Reference number: S54817  
A/Acession: S54817  
A/Status: preliminary  
A/Molecule type: mRNA  
A/Residues: 1-177 <NC>  
A/Cross-references: EMBL:X84237, NID:g809045, PIDN:CMA59015.1, PID:g809046  
Keywords: T-cell receptor

Query Match	24.8%;	Score 605;	DB 2;	Length 177;
Best Local Similarity	99.2%;	Pred. No. 1.2e-34;		
Matches 117; Conservative	0;	Mismatches 1;	Indels 0;	Gaps 0

QY	Db	QY
333 LEDPGLCYLDDGILFIYGVIIITALLCPAKRSRAEYNAALOPDNOYLNLNLNGREEEYV	382	333 LEDPGLCYLDDGILFIYGVIIITALLCPAKRSRAEYNAALOPDNOYLNLNLNGREEEYV
26 LDDPLCYLDDGILFIYGVIIITALLYLAKRSRAEYNAALOPDNOYLNLNLNGREEEYV	85	26 LDDPLCYLDDGILFIYGVIIITALLYLAKRSRAEYNAALOPDNOYLNLNLNGREEEYV
353 LEKRRADPEWGGQQRNRPDQEVYTNALQDMMALAYSEIGTKERRRRGKHDDGLYQ	440	353 LEKRRADPEWGGQQRNRPDQEVYTNALQDMMALAYSEIGTKERRRRGKHDDGLYQ
86 LEKRRADPEWGGQQRNRPDQEVYTNALQDMMALAYSEIGTKERRRRGKHDDGLYQ	143	86 LEKRRADPEWGGQQRNRPDQEVYTNALQDMMALAYSEIGTKERRRRGKHDDGLYQ

RESULT 6  
A35900  
T-cell receptor CD3 eta chain precursor - mouse



C:Species: Mus musculus (house mouse)  
C:Date: 23-Oct-1990 #sequence\_revision 23-Oct-1990 #text\_change 09-Jul-2004  
C:Accession: A35900; A60374; G46522  
R:Jin, Y.J.; Clayton, L.K.; Howard, F.D.; Koyasu, S.; Steh, M.; Steindrich, R.; Tarr, G.  
Proc. Natl. Acad. Sci. U.S.A. 87, 3319-3323, 1990  
A:Title: Molecular cloning of the CD3eta subunit identifies a CD3eta-related product in  
A:Reference number: A35900; MUID:90239005; PMID:2139725  
A:Accession: A35900  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-206 <JIN>  
A:Cross-references: UNIPROT:P29020; GB:M3158; NID:G192488; PIDN:AAA37398.1; PID:G309159  
R:Ohno, H.; Saito, T.  
Int. Immunol. 2, 1117-1119, 1990  
A:Title: CD3zeta and eta chains are produced by alternative splicing from a common gene.  
A:Reference number: A60374; MUID:91190781; PMID:2150596  
A:Accession: A60374  
A:Status: not compared with conceptual translation  
A:Molecule type: DNA  
A:Residues: 144-206 <OHNO>  
R:Jensen, J.P.; Cenciarelli, C.; Hou, D.; Rellahan, B.L.; Dean, M.; Weisman, A.M.  
J. Immunol. 150, 122-130, 1993  
A:Title: T cell antigen receptor-eta subunit. Low levels of expression and limited cross  
A:Reference number: A46522; MUID:93107707; PMID:8417118  
A:Contents: annotation  
C:Comment: The functional significance of this alternatively spliced product of the CD3  
lno acids, differs widely among various mammalian species in sequence, length, and even  
C:Keywords: alternative splicing; T-cell receptor; transmembrane protein

Query Match 24.8%; Score 605; DB 2; Length 206;  
Best Local Similarity 99.2%; Pred. No. 1.4e-34;  
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

OY 323 LEDPRLCYLIDGILFIYVITITALLYLRKFSRSASATAANLDPNQLYNELNGRREEDV 382  
DB 26 LLDPLKCYLIDGILFIYVITITALLYLRKFSRSASATAANLDPNQLYNELNGRREEDV 85

OY 383 LEKRRAPDEMGKQORRRNPOEGVYNALQKDKMAEAYSEIGTGERRRGKHDLQ 440  
DB 86 LEKRRAPDEMGKQORRRNPOEGVYNALQKDKMAEAYSEIGTGERRRGKHDLQ 143

RESULT 7  
A31768  
T-cell receptor zeta chain precursor - human  
C:Species: Homo sapiens (man)  
C:Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
C:Accession: A31768  
R:Weisman, A.M.; Hou, D.; Orloff, D.G.; Modi, W.S.; Sevanetz, H.; O'Brien, S.J.; Klausne  
Proc. Natl. Acad. Sci. U.S.A. 85, 9709-9713, 1988  
A:Title: Molecular cloning and chromosomal localization of the human T-cell receptor zet  
A:Reference number: A31768; MUID:89071765; PMID:2974162  
A:Accession: A31768  
A:Molecule type: mRNA  
A:Residues: 1-163 <WEI>  
A:Cross-references: UNIPROT:P20963; GB:J04132; NID:G623041; PIDN:AAA60394.1; PID:G623042  
C:Keywords: phosphoprotein; T-cell receptor; transmembrane protein  
F:1-21/Domain: signal sequence #status predicted <SIG>  
F:22-163/Product: T-cell receptor zeta chain #status predicted <MAT>

Query Match 24.8%; Score 604.5; DB 2; Length 163;  
Best Local Similarity 84.9%; Pred. No. 1.2e-34;  
Matches 118; Conservative 6; Mismatches 14; Indels 1; Gaps 1;

OY 323 LEDPRLCYLIDGILFIYVITITALLYLRKFSRSASATAANLDPNQLYNELNGRREEDV 382  
DB 26 LLDPLKCYLIDGILFIYVITITALLYLRKFSRSASATAANLDPNQLYNELNGRREEDV 85

OY 383 LEKRRAPDEMGKQORRRNPOEGVYNALQKDKMAEAYSEIGTGERRRGKHDLQ 442  
DB 86 LDKRRGRDPEMGK-QRRKNPQEGLYNELQKDKMAEAYSEIGTGERRRGKHDLQ 144

OY 443 STATKDTYDALHMQALPR 461

DB 145 STATKDTYDALHMQALPR 163  
|||||  
RESULT 8  
A45089  
T-cell receptor CD3 theta chain, alternate splice form - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 05-Nov-1999  
C:Accession: A45089; I49587  
R:Clayton, L.K.; Diener, A.C.; Lerner, A.; Tee, A.G.; Koyasu, S.; Reinherz, E.L.  
J. Biol. Chem. 267, 26023-26030, 1992  
A:Title: Differential regulation of T-cell receptor processing and surface expression af  
A:Reference number: A45089; MUID:93100325; PMID:1464613  
A:Accession: A45089  
A:Status: preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-188 <CLA>  
A:Cross-references: GB:S51932; NID:G261998; PIDN:AAB24559.1; PID:G261999  
A:Experimental source: thymus  
A:Note: Sequence extracted from NCBI backbone (NCBI:P120865)  
R:Lerner, A.; Diener, A.C.; Reinherz, E.L.; Clayton, L.K.  
Eur. J. Immunol. 22, 2135-2140, 1992  
A:Title: Human genomic sequences corresponding to murine CD3eta-related transcripts: lacc  
A:Reference number: I49587; MUID:92347411; PMID:1322304  
A:Accession: I49587  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-188 <RES>  
A:Cross-references: GB:L03353; NID:G192508; PIDN:AAA37401.1; PID:G192509  
C:Keywords: T-cell receptor

Query Match 24.1%; Score 589.5; DB 2; Length 188;  
Best Local Similarity 98.3%; Pred. No. 1.5e-33;  
Matches 116; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

OY 323 LEDPRLCYLIDGILFIYVITITALLYLRKFSRSASATAANLDPNQLYNELNGRREEDV 382  
DB 26 LLDPLKCYLIDGILFIYVITITALLYLRKFSRSASATAANLDPNQLYNELNGRREEDV 85

OY 383 LEKRRAPDEMGKQORRRNPOEGVYNALQKDKMAEAYSEIGTGERRRGKHDLQ 440  
DB 86 LEKRRAPDEMGK-QRRKNPQEGVYNALQKDKMAEAYSEIGTGERRRGKHDLQ 142

RESULT 9  
S35759  
BHD9D10 protein - mouse (fragment)  
C:Species: Mus musculus (house mouse)  
C:Date: 13-Jan-1995 #sequence\_revision 13-Jan-1995 #text\_change 21-Jan-2000  
C:Accession: S35759  
R:Froyen, G.F.V.  
Submitted to the EMBL Data Library, May 1993  
A:Reference number: S35759  
A:Accession: S35759  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-136 <FRO>  
A:Cross-references: EMBL:X72796; NID:G312496; PIDN:CAAS1316.1; PID:G312497  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 24.1%; Score 588.5; DB 2; Length 136;  
Best Local Similarity 82.6%; Pred. No. 1.2e-33;  
Matches 114; Conservative 7; Mismatches 14; Indels 3; Gaps 2;

OY 1 MAWVTLFLTAAAK-VPKQIQVQSGPELKKPGETYKISCKAGYEPFTYGMWVQAP 59  
DB 1 MAWVTLFLTAAAGSQAQIQVQSGPELKKPGETYKISCKAGYFTTDGMWVQAP 60

OY 60 GQGLKMGWINTSTGESFPADDFGRDFSLTSSANTAYLIQINLKEDMATYFCARWV 119  
DB 61 GQGLKMGWINTYGESYVDDFGRDFVSLTSSASAAVLIQINLKEDMATYFCARGF 120

OY 120 YHGVPYWGQGTIVTSS 137  
| : ||||| :  
DB 121 Y--AMDYWGQGTIVTSS 136

## RESULT 10

S33905  
Ig heavy chain precursor V region - synthetic

C:Species: synthetic  
C:Date: 13-Jan-1995 #sequence\_revision 30-Apr-1998 #text\_change 20-Oct-2000  
C:Accession: S33905  
R:Lin, A.Y.; Robinson, R.R.; Hellstrom, K.E.; Murray Jr., E.D.; Chang, C.P.; Hellstrom  
Proc. Natl. Acad. Sci. U.S.A. 84, 3439-3443, 1987  
A:Title: Chimeric mouse-human IgG1 antibody that can mediate lysis of cancer cells.  
A:Reference number: S33905; MUID:87204152; PMID:3106970  
A:Accession: S33905  
A:Molecule type: mRNA  
A:Residues: 1-146 <L10>

A:Cross-references: EMBL:M16072; NID:g195270; PIDN:AAA38229.1; PID:g195271

Query Match 24.0%; Score 585; DB 4; Length 146;  
Best Local Similarity 80.6%; Pred. No. 2,2e-33;  
Matches 112; Conservative 9; Mismatches 16; Indels 2; Gaps 2;

OY 1 MAMVWTLFLMAAAK-VPKQIQLVQSGPELKKPGETVKISCKASGYPTNYGMWVKQAP 59  
| : ||||| :  
DB 1 MDMLNLLFLMAAAQSAQAQIQLVQSGPELKKPGETVKISCKASGYPTNYGMWVKQAP 60

OY 60 GQGLKMWGINTSGESTFADDFKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEV 119  
| : ||||| :  
DB 61 GKGLKMWGINTYGTQPTVADDFKGRFAFSLETSAYTAYLQINNLKNEDEMATYFCARSY 120

OY 120 YHG--VYVYWGQGTIVTSS 137  
| : ||||| :  
DB 121 GNSRISDYWGQGTIVTSS 139

## RESULT 11

PH1225  
Ig heavy chain precursor V region (M-1151) - mouse (fragment)

C:Species: Mus musculus (house mouse)  
C:Date: 17-Apr-1993 #sequence\_revision 17-Apr-1993 #text\_change 21-Jan-2000  
C:Accession: PH1225  
R:Weissenborn, W.; Scheuer, W.; Kaluza, B.; Schwitzke, M.; Reiter, C.; Flieger, D.; Lenz  
Gene 121, 271-278, 1992  
A:Title: Combinatorial functions of two chimeric antibodies directed to human CD4 and or  
A:Reference number: PH1224; MUID:93077041; PMID:1446824  
A:Accession: PH1225  
A:Molecule type: mRNA  
A:Residues: 1-139 <WE1>

A:Cross-references: GB:S50263; NID:g260763; PIDN:AA824319.1; PID:g260764  
A:Note: this mouse sequence was hybridized and fused with a human constant region gene  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:1-19/Domain: signal sequence #status predicted <SIG>  
F:20-139/Product: Ig heavy chain V region #status predicted <MAT>  
F:34-117/Domain: immunoglobulin homology <IMM>

Query Match 22.7%; Score 555; DB 2; Length 139;  
Best Local Similarity 76.3%; Pred. No. 2,4e-31;  
Matches 106; Conservative 13; Mismatches 18; Indels 2; Gaps 2;

OY 1 MAMVWTLFLMAAAK-VPKQIQLVQSGPELKKPGETVKISCKASGYPTNYGMWVKQAP 59  
| : ||||| :  
DB 1 MAMVWTLFLMAAAQSTQAOIQLVQSGPELKKPGETVKISCKASGYPTNYGMWVKQAP 60

OY 60 GQGLKMWGINTSGESTFADDFKGRFDSLETSANTAYLQINNLKSEDMATYFCARME 118  
| : ||||| :  
DB 61 GKGLKMWGINTYGTQPTVADDFKGRFAFSLETSAYTAYLQINNLKNEDEMATYFCARHY 120

OY 119 YHGVPYWGQGTIVTSS 137  
| : ||||| :  
DB 121 Y--AMDYWGQGTIVTSS 136

DB 121 AYGDPLDYWGQGTIVTSS 139

## RESULT 12

JC4664  
T-cell receptor zeta chain - bovine

C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #text\_change 05-Nov-1999  
C:Accession: JC4664  
R:Hagens, G.; Galley, Y.; Glaser, I.; Davis, W.C.; Baldwin, C.L.; Clevers, H.; Dobbelaert  
Gene 169, 165-171, 1996  
A:Title: Cloning, sequencing and expression of the bovine CD3 epsilon and TCR-zeta chains  
A:Reference number: JC4663; MUID:96194796; PMID:8647441  
A:Accession: JC4664  
A:Molecule type: mRNA  
A:Residues: 1-166 <HAG>

A:Cross-references: GB:U25688; NID:g1263011; PIDN:AA48548.1; PID:g1263012  
C:Comment: This protein plays a pivotal role in linking T-cell receptor-triggering to sev  
lymphokine receptor gene expression.  
C:Genetics:

A:Gene: tcr-zeta  
C:Keywords: GTP binding; signal transduction; T-cell receptor  
F:129-146/Region: GDP/GTP-binding

Query Match 22.1%; Score 539; DB 2; Length 166;  
Best Local Similarity 75.2%; Pred. No. 3,7e-30;  
Matches 106; Conservative 11; Mismatches 22; Indels 2; Gaps 1;

OY 323 LEDPKLCYLLDGLIFVYVITLALYLRKFSRSSETAANLQDPNOLYNELNGRREYDV 382  
| : ||||| :  
DB 26 LDPKLCYLLDGLIFVYVITLALFLRAKFSRSANNPAYQGGNPYNELNVRREYAV 85

OY 383 LEKRRADPEMGKQQRNRRPQEGYVNAQKDMAEAYSIEIGTK--GERRRGKHGLYQ 440  
| : ||||| :  
DB 86 LDRRGDFPEKGGKPKRKNPNVEYNELRKDMAEAYSIEIGKSDNRRRGKHGLYQ 145

OY 441 GLSTATKDYTDALHMOTLAPR 461  
| : ||||| :  
DB 146 GLSTATKDYTDALHMOTLAPR 166

## RESULT 13

I46424  
T-cell surface glycoprotein CD3 zeta chain - sheep

C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)  
C:Date: 16-Aug-1996 #sequence\_revision 16-Aug-1996 #text\_change 03-Jul-2004  
C:Accession: I46424; S22980  
R:Hein, W.R.; Tunacliffe, A.  
Immunogenetics 37, 279-284, 1993  
A:Title: Invariant components of the sheep T-cell antigen receptor: cloning of the CD3 ei  
A:Reference number: I46424; MUID:9311305; PMID:8420837  
A:Accession: I46424  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-166 <HE1>

A:Cross-references: UNIPROT:P29329; EMBL:Z12968; NID:g1399; PIDN:CAA78312.1; PID:g1400  
C:Keywords: glycoprotein

Query Match 22.1%; Score 539; DB 2; Length 166;  
Best Local Similarity 74.5%; Pred. No. 3,7e-30;  
Matches 105; Conservative 13; Mismatches 21; Indels 2; Gaps 1;

OY 323 LEDPKLCYLLDGLIFVYVITLALYLRKFSRSSETAANLQDPNOLYNELNGRREYDV 382  
| : ||||| :  
DB 26 LDPKLCYLLDGLIFVYVITLALFLRAKFSRSADNPAYQGGNPYNELNVRREYAV 85

OY 383 LEKRRADPEMGKQQRNRRPQEGYVNAQKDMAEAYSIEIGTK--GERRRGKHGLYQ 440  
| : ||||| :  
DB 86 LDRRGDFPEMGKQQRKKNPNVEYNELRKDMAEAYSIEIGKSDNRRRGKHGLYQ 145

OY 441 GLSTATKDYTDALHMOTLAPR 461  
| : ||||| :  
DB 146 GLSTATKDYTDALHMOTLAPR 166

## RESULT 14

B42848  
I6 mAb heavy chain V region - mouse (fragment)

C:Species: Mus musculus (house mouse)  
C:Date: 27-Apr-1993 #sequence\_revision 18-Nov-1994 #text\_change 21-Jan-2000  
C:Accession: B42848; S33903  
R: Fell, H.P.; Gayle, M.A.; Yelton, D.; Lipsich, L.; Schieven, G.L.; Marken, J.S.; Aruffo  
J. Biol. Chem. 267, 15552-15558, 1992  
A:Title: Chimeric I6 anti-tumor antibody. Genomic construction, expression, and character  
A:Reference number: A42848; MUID:92348410; PMID:1639794  
A:Accession: B42848  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-120 <FE1>  
A:Cross-references: GB:M90690; NID:g195065; PTD:AAA8146.1; PTD:g195066  
A>Note: Sequence extracted from NCBI backbone (NCBIN:109960, NCBI:P:109961)  
A:Accession: S33903  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-120 <FE2>  
A:Cross-references: EMBL:M90691  
C:Superfamily: Immunoglobulin V region; immunoglobulin homology  
F:15-98/Domain: immunoglobulin homology <IMM>

Query Match 22.0%; Score 537.5; DB 2; Length 120;  
Best Local Similarity 84.2%; Pred. No. 3.2e-30;

Matches 101; Conservative 7; Mismatches 11; Indels 1; Gaps 1;

QY 19 QIQLVQSGPELKKPGETVKISCKASGYPTNYGMNWKQAPQGGLKMMGWINSTGSTF 78

Db 1 QIQLVQSGPELKKPGETVKISCKASGYPTNYGMNWKQAPQGGLKMMGWINSTGSTF 60

QY 79 ADDKGRDFSLFSTANTAYLQINNLSKEDMATYFCARWEVHG-YVPYWGQSTIVTSS 137

Db 61 ADDKGRDFSLFSTANTAYLQINNLSKEDMATYFCARWEVHG-YVPYWGQSTIVTSS 120

## RESULT 15

PH1404  
Ig heavy chain V region (clone micro m+ 46-12) - mouse (fragment)

C:Species: Mus musculus (house mouse)  
C:Date: 31-Dec-1993 #sequence\_revision 31-Dec-1993 #text\_change 21-Jan-2000  
C:Accession: PH1404; PH1406  
R: Shirasawa, T.; Miyazoe, I.; Higawara, S.; Kimoto, H.; Shigemoto, K.; Taniguchi, M.; Ta  
J. Exp. Med. 176, 1209-1214, 1992  
A:Title: Heavy chain variable (VH) region diversity generated by VH gene replacement in  
la virus.  
A:Reference number: PH1403; MUID:93018837; PMID:1402663  
A:Accession: PH1404  
A:Molecule type: DNA  
A:Residues: 1-124 <SH1>  
A:Accession: PH1406  
A:Molecule type: DNA  
A:Residues: 115-121 <SH2>  
C:Genetics:  
A:Insertions: 16/1  
C:Superfamily: immunoglobulin V region; immunoglobulin homology  
C:Keywords: heterotetramer; immunoglobulin  
F:35-118/Domain: immunoglobulin homology <IMM>

Query Match 22.0%; Score 536; DB 2; Length 124;  
Best Local Similarity 85.6%; Pred. No. 4.2e-30;

Matches 101; Conservative 6; Mismatches 9; Indels 2; Gaps 1;

QY 1 MAMVTTLFLMAAAK--VPKQIQLVQSGPELKKPGETVKISCKASGYPTNYGMNWKQ 58

Db 1 MDMLKNNLFLMAAAQTGAQIQLVQSGPELKKPGETVKISCKASGYPTNYGMNWKQ 60

QY 59 PQGGLKMMGWINSTGSTFADDFKGRDFSLFSTANTAYLQINNLSKEDMATYFCAR 116

Db 61 PQGGLKMMGWINSTGSTFADDFKGRDFSLFSTANTAYLQINNLSKEDMATYFCAR 118

Search completed: April 20, 2005, 07:07:48  
Job time : 18.8485 secs







Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)

cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC Building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.

Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

## FEATURES

## Source

Location/Qualifiers  
1..953  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="taxon:10090"  
/clone="E430003L03"  
/issue\_type="thymus"  
/cell\_type="thymic cells"  
/clone\_id="RIKEN full-length enriched, 2 days neonate thymus thymic cells (NOD)"

## ORIGIN

## Alignment Scores:

Pred. No.: 1.28e-60 Length: 953  
Score: 715.00 Matches: 146  
Percent Similarity: 86.39% Conservative: 0  
Best Local Similarity: 86.39% Mismatches: 11  
Query Match: 29.29% Indels: 12  
DB: 6 Gaps: 2

US-09-596-774-6 (1-461) x BY746048 (1-953)

QY 293 SerProValHisProThrGlyThrSerGlnProGlnArgProGluAspCysArgProArg 312  
DB 94 TCGCTGTCATCTCCACGTCGCTCCACGAGCAG-----AGCAGCAGA 138  
QY 313 GlySerValIlybGlyThrGlyLeuAspPheLeuGluAspProIyLeuCysTyrLeuLeu 332  
DB 139 GCTTTGGCTCG-----CTGGATCCCAAACTGTGCTACTGTGCTA 177  
QY 333 AspGlyIleuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaIlybPhe 352  
DB 178 GATGGAATCCTCTTCATCTACGAGTCATCATCACAGCCCTGTACTGAGAGCAAAATTC 237  
QY 353 SerArgSerAlaGluThrAlaAlaAsnLeuGlnAspProArgGlnIleuTyrAsnGluLeu 372  
DB 238 AGCAGAGGTGAGAGACTGCTGCCAACTGACGAGCCCAACCACTCTACAAATGAGCTC 297  
QY 373 AsnLeuGlyArgArgGluGluTyrAspValIleuGluIlybIlybArgAlaArgAspProGlu 392  
DB 298 AATCTAGGCGCAAGAGAGGATATGACTCTTGAGAAAGACGGGCTCCGGATCCAGAG 357  
QY 393 MetGlyIlybGlnGlnArgArgArgAsnProGlnGlnIlybValTyrAsnAlaLeuGln 412  
DB 358 ATGGAGGCAAAACAGCAGAGAGAGAGAAACCCCGAAGCGGTATCAATGACTGCTGAG 417  
QY 413 IlybAspIlybMetAlaGluAlaTyrSerGluIleGlyThrIlybGlyIlybValArgArgGly 432  
DB 418 AAAACAGAGATGAG 477  
QY 433 IlybGlyHisAspGlyLeuTyrGlnIlybLeuSerThrAlaThrIlybAspThrTyrAspAla 452  
DB 478 AAGGGGCGCAGATGGCTTTACAGAGGTCTTCAGCACTGCCACCAAGAGACCTATGATGCC 537  
QY 453 LeuHisMetGlnThrLeuAlaProArg 461  
DB 538 CTGCATATGACAGACCCCTGGCCCTGCG 564

## RESULT 3

## AC088087

## LOCUS

## DEFINITION

AK088087 1570 bp. mRNA. linear. HTC 03-APR-2004  
Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN  
full-length enriched library, clone: E430003L03 product: CD3 antigen,  
zeta polypeptide, full insert sequence.

## ACCESSION

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## MEDLINE

## PUBMED

## REFERENCE

## AUTHORS

## TITLE

1 Carninci, P. and Hayashizaki, Y.  
High-efficiency full-length cDNA cloning  
Meth. Enzymol. 303, 19-44 (1999)

2 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,  
Itoh, M., Kono, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.  
Normalization and subtraction of cap-trapper-selected cDNAs to  
prepare full-length cDNA libraries for rapid discovery of new genes  
Genome Res. 10 (10), 1617-1630 (2000)

3 Shibata, K., Itoh, M., Aizawa, K., Nagao, S., Sasaki, N., Carninci, P.,  
Kono, H., Akiyama, J., Nishi, K., Kitahara, T., Tashiro, H., Itoh, M.,  
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A.,  
Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kasai, Y.,  
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matsubara, S.,  
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J.,  
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.  
RIKEN integrated sequence analysis (RISA) system-384-format  
sequencing pipeline with 384 multiplexed sequencer  
Genome Res. 10 (11), 1757-1771 (2000)

4 The RIKEN Genome Exploration Research Group Phase II Team and the  
FANTOM Consortium.  
Functional annotation of a full-length mouse cDNA collection  
Nature 409, 685-690 (2001)

5 The FANTOM Consortium and the RIKEN Genome Exploration Research  
Group Phase I & II Team.  
Analysis of the mouse transcriptome based on functional annotation  
of 60,770 full-length cDNAs  
Nature 420, 563-573 (2002)

6 (bases 1 to 1570)  
Adachi, J., Aizawa, K., Akimura, T., Arai, A., Hashizume, W.,  
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,  
Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T.,  
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,  
Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M.,  
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,  
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,  
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,  
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,  
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S.,  
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A.,  
Muramatsu, M. and Hayashizaki, Y.

Direct Submission  
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of  
Physical and Chemical Research (RIKEN), Laboratory for Genome  
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),  
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,  
Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,  
URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222,  
Fax: 81-45-503-9216)  
cDNA library was prepared and sequenced in Mouse Genome

## COMMENT

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.

Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
Trust/MRC Building Addenbrookes Hospital Cambridge) whose  
assistance we gratefully acknowledge.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.jp/>  
URL: <http://fantom.gsc.riken.jp/>  
Location/Qualifiers

## FEATURES

SOURCE

1..1570  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="FANTOM\_DB:E430003L03"  
/db\_xref="taxon:10090"  
/clone="E430003L03"  
/cell\_type="thymic cells"  
/tissue\_type="thymus"  
/clone\_id="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="2 days neonate"  
76..571  
/note="CD3 antigen, zeta polypeptide (MGD|MG:88334,  
GB|NM\_031162, evidence: BLASTN, 100%, match=536)  
putative"  
1549..1554  
/note="putative"  
1570  
/note="putative"  
polya\_signal  
polya\_site  
ORIGIN

## Alignment Scores:

Pred. No.: 2,6e-60 Length: 1570  
Score: 715.00 Matches: 146  
Percent Similarity: 86.39% Conservative: 0  
Best Local Similarity: 86.39% Mismatches: 11  
Query Match: 29.29% Indels: 12  
DB: 3 Gaps: 2

US-09-596-774-6 (1-461) x AK088087 (1-1570)

QY 293 SerProValHisProThrGlyThrSerGlnProGlnArgProGlnAspGlyArgProArg 312  
DB 98 TCGCGTGCATCTCCACGTCGCGGTTCCAGAGAGAG-----AGGCACAGA 142  
QY 313 GlySerValValGlyThrGlyLeuAspPheLeuGlnAspProGlyLeuGlyLeu 332  
DB 143 GCTTGGTCTCG-----CTGCATCCCAAACTCTGCTACTGCTTA 181  
QY 333 AspGlyLeuPheLeuGlyThrGlyValAlaIleThrAlaLeuTyrlleuAlaGlyPhe 352  
DB 182 GATGGAATCCTCTTCATCTACGGAGTCATATCAACCCCTGTAACGAGAGCAAAATTC 241  
QY 353 SerArgSerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrlleuGlu 372  
DB 242 AGCAGAGATGAGAGACTGCTGCCAACCTCCAGACCCCAACAGCTTACATAGAGCTC 301  
QY 373 AsnLeuGlyArgArgGluGluTyrlleuValLeuGlyLeuArgAlaArgAspProGlu 392  
DB 302 AATCTAGAGCGAAGAGAGGAATATGACGTTTGGAAAGAGAGGGGCTGGGATCCAGAG 361  
QY 393 MetGlyGlyArgGlnGlnArgArgGlnArgGlnArgGlnArgGlnArgGlnArgGln 412  
DB 362 ATGGAGAGCAAAAG 421  
QY 413 LysAspLysMetAlaGluAlaTyrlleuGlyValIleGlyThrGlyGlyLeuArgArgGly 432  
DB 422 AAAGACAAAGTGCACAAAGCTTACATGATCGGACAAAGAGGAGAGAGAGAGAGAG 481  
QY 433 LysGlyHisAspGlyLeuTyrlleuGlnLeuSerThrAlaThrLysAspThrTyrlleu 452

DB 482 AAGGGGACAGATGCGCTTTACAGGGTCTCAGCACTGCCACCAAGAGACACTATGATGCC 541  
QY 453 LeuHisMetGlnThrLeuAlaProArg 461  
DB 542 CTGCATATGACAGACCTTGGCCCTCGC 568

RESULT 4  
CA577711  
LOCUS  
DEFINITION  
CA577711 577 bp mRNA linear EST 19-NOV-2002  
CDNA library (Long) Mus musculus cDNA clone NIA:K0713C05  
IMAGE:30074812 5', mRNA sequence.

ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
Mus musculus  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
1 (bases 1 to 577)  
Piao, Y., Kargul, G.J., Dudekula, D.B., Qian, Y., Luo, A., Carter, M.G.,  
Alba, K., Taub, D., Longo, D.L., Keller, J. and Ko, M.S.H.  
Systematic Analyses of NIA Mouse Hematopoietic Stem Cell  
(lin-/c-Kit-/Sca-1-) cDNA library (Long)  
Unpublished (2001)  
Other ESTs: K0713C05-3  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: [cdna@igsun.grc.nia.nih.gov](mailto:cdna@igsun.grc.nia.nih.gov)  
Plate: K0713 Row: C Column: 05  
Seq primer: M13 Reverse  
High quality sequence stop: 577  
POLYA=No.

## FEATURES

SOURCE

1..577  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6NCR"  
/db\_xref="nlaEST:K0713C05-5N"  
/db\_xref="taxon:10090"  
/clone="NIA:K0713C05 IMAGE:30074812"  
/tissue\_type="Hematopoietic Stem Cell  
(lin-/c-Kit-/Sca-1-)"  
/dev\_stage="Age approx. 10 weeks old"  
/lab\_host="DH10B"  
/clone\_id="NIA Mouse Hematopoietic Stem Cell  
(lin-/c-Kit-/Sca-1-) cDNA library (Long)"  
/note="Vector: pSPORT1 (Invitrogen); Site 1: SalI; Site 2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (<http://igsun.grc.nia.nih.gov/cDNA>). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11: 1553-1558 (2001)). [PMID: 11544199]. Total RNAs were  
obtained from Drs. Dennis Taub, Dan Longo (National  
Institute on Aging, USA), Jonathan Keller (National Cancer  
Institute, USA). Double-stranded cDNAs were synthesized  
with an Oligo(dT) primer (Invitrogen:  
5'-pGACTAGTCTAGATCGCAGCGGCGGCTTTT-3') from  
0.9 ug of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to Lome-Linker lin-Sal4, purified by phenol/chloroform, and  
separated from free linkers by centrifugation. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
Ex Tag polymerase (Takara) with a primer Sal4-S. The  
products were purified by phenol/chloroform and centrifuged  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pSPORT1 plasmid vector.  
The DH10B E. coli host was transformed with the ligation  
mixture by the standard chemical method. The average  
insert size is about 2.1 kb. The library was constructed  
by Yulan Piao (NIA)."



ORIGIN	Alignment Scores:	Pred. No.:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
US-09-596-774-6 (1-461) x CA577711 (1-577)	9.88e-61	Length: 577	713.00	99.28%	Mismatches: 0	Indels: 0	Gaps: 0
QY	LeuGluAaPProLyLeuCYeTYrLeuLeuAaPgiYiLeuPheIeTYrGlyValIle	342					
Db	133 CTGCTGAGATCCCAAACTCTGCTACTTGTAGATGAAATCCTTTATCTACGAGCTCATC	192					
QY	343 ILeThraLaLeuTYrLeuAaRgAlaLySbPheSeRaGSeRaIaGluThRaLaLaAaLeu	362					
Db	193 ATCAGAGCCCTGTACTCTAGAGCAAAATTCAGCAGAGATGAGAGACTGTGCTCCAACTTG	252					
QY	363 GlnAaPProAaEngInLeuTYrAaEngIuLeuAaLeuGlyAaGAGGluTYrAaPAl	382					
Db	253 CAGAGCCCAACCAAGCTCTTACATAGCTCAATCTAGGCGCAAGAGAAATATGACCTC	312					
QY	383 LeuGluLybLYbAaRgAlaAaPProGluMetGlyGlyLYbGlnGlnAaAaRgAa	402					
Db	313 TTGGAAAGAAAGCGGCTCGGATCCAGAGATGGAGGCAACACAGAGAGAGAGAAC	372					
QY	403 ProGlnGluGlyValTYrAaSnAlaLeuGlnLybAaPlybMeaAlaGluAlaTYrSeRglu	422					
Db	373 CCCAGAGAGCGCTATACATAGCACTGCAGAAAGCAAGATGGCAGAAAGCTACAGTAG	432					
QY	423 ILeGlyThrLybGlyGlnAaRgAaRgGlyLybGlybIbAaPglyLeuTYrGlnGlyLeu	442					
Db	433 ATCGGCAAAAGCGAGAGCGGAGAGGCAAGGCGGCAAGATGGCTTTACAGAGGCTTC	492					
QY	443 SerThraIaThrLybAaPThrTYrAaSnAlaLeuHISmetGlnThThraLaAaProAaRg	461					
Db	493 AGCACTGCCACCAAGACACCTATGATGCTCCCTGCATATGACAGACCCTGACCCCTGCG	549					
RESULT 5							
AKO17904		1598 bp	mRNA	linear	HTC 03-APR-2004		
LOCUS							
DEFINITION	Mus musculus adult male thymus cDNA, RIKEN full-length enriched library, clone:583047D18 product:CD3 antigen, zeta polypeptide, full insert sequence.						
ACCESSION	AKO17904						
VERSION	AKO17904.1						
KEYWORDS	HTC; CAP trapper.						
SOURCE	Mus musculus (house mouse)						
ORGANISM	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.						
REFERENCE							
AUTHORS	1 Carninci, P. and Hayashizaki, Y						
TITLE	High-efficiency full-length cDNA cloning						
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)						
MEDLINE	99279253						
PUBMED	10349636						
REFERENCE	2						
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.						
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes						
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)						
MEDLINE	20499374						
PUBMED	11042159						
REFERENCE	3						
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsuana, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujitake, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wataniki, M.,						

JOURNAL	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
MEDLINE	RIKEN integrated sequence analysis (RISA) system--384-format
PUBMED	sequencing pipeline with 384 multicapillary sequencer
REFERENCE	Genome Res. 10 (11), 1757-1771 (2000)
AUTHORS	11076861
TITLE	4
JOURNAL	The RIKEN Genome Exploration Research Group Phase II Team and the
MEDLINE	FANTOM Consortium.
PUBMED	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	5
TITLE	The FANTOM Consortium and the RIKEN Genome Exploration Research
JOURNAL	Group Phase I & II Team.
MEDLINE	Analysis of the mouse transcriptome based on functional annotation
PUBMED	of 60,770 full-length cDNAs
REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	6 (bases 1 to 1598)
TITLE	Adachi, J., Aizawa, K., Akhira, S., Akimura, T., Arai, A., Aono, H.,
JOURNAL	Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y.,
MEDLINE	Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K.,
PUBMED	Hiraoaka, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M.,
REFERENCE	Kesukawa, T., Kato, H., Kawai, J., Kojima, Y., Konno, H., Kouda, M.,
AUTHORS	Komura, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K.,
TITLE	Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C.,
JOURNAL	Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D.,
MEDLINE	Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y.,
PUBMED	Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Tanaka, T.,
REFERENCE	Tajima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K.,
AUTHORS	Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-VUL-2000) Yoshinide Hayashizaki, The Institute of
MEDLINE	Physical and Chemical Research (RIKEN), Laboratory for Genome
PUBMED	Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
REFERENCE	RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
AUTHORS	Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp,
TITLE	URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222,
JOURNAL	fax: 81-45-503-9216)
MEDLINE	Please visit our web site (http://genome.gsc.riken.jp/) for further
PUBMED	details.
REFERENCE	cDNA library was prepared and sequenced in Mouse Genome
AUTHORS	Encyclopedia Project of Genome Exploration Research Group in Riken
TITLE	Genomic Sciences Center and Genome Science Laboratory in RIKEN.
JOURNAL	Division of Experimental Animal Research in Riken contributed to
MEDLINE	prepare mouse tissues. First strand cDNA was primed with a primer
PUBMED	[5', GAGAGGAGGATCGAGTATTAATTAATATCCGCCGCCGCC 3']. cDNA was
REFERENCE	prepared by using retrovirus thermo-activated reverse transcriptase
AUTHORS	and subsequently enriched for full-length by cap-trapper. cDNA went
TITLE	through one round of normalization to Rot = 10.0 and subtraction to
JOURNAL	Rot = 100.0. Second strand cDNA was prepared with the primer
MEDLINE	adaptor of sequence [5'
PUBMED	GAGAGGAGGATCGAGTATTAATTAATATCCGCCGCCGCC 3']. cDNA was cleaved
REFERENCE	with BamHI and XhoI. Vector: a modified plasmid KS(+) after
AUTHORS	bulk excision from lambda FLC I. Cloning sites, 5' end: SalI, 3'
TITLE	end: BamHI. Host: DH10B.
JOURNAL	Location/Qualifiers
MEDLINE	1. 1598
PUBMED	/organism="Mus musculus"
REFERENCE	/mol_type="mRNA"
AUTHORS	/strain="C57BL/6J"
TITLE	/db_xref="FANTOM,DB:5830407D18"
JOURNAL	/id_xref="taxon:10090"
MEDLINE	/clone="5830407D18"
PUBMED	/sex="male"
REFERENCE	/tissue_type="thymus"
AUTHORS	/clone_lib="RIKEN full-length enriched mouse cDNA library"
TITLE	/dev_stage="adult"
JOURNAL	112. 606
MEDLINE	/note="unnamed protein product; CD3 antigen, zeta
PUBMED	polypeptide (MGI:88334, GB NM_031162, evidence:
REFERENCE	BLASTN, 100%, match=536)
AUTHORS	putative"



IMAGE:7384006 5', mRNA sequence.  
 ACCESSION COS68334  
 VERSION COS68334.1 GI:50380963  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 764)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics / NIH  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical College of Wisconsin  
 CDNA Library Preparation: Open Biosystems  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLM15548 row: d column: 20  
 High quality sequence stop: 746.  
 Location/Qualifiers  
 1..764  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="IMAGE:7384006"  
 /lab\_host="DH10B Tona"  
 /clone\_1lb="NIH\_MGC\_250"  
 /note="Organ: thymus; Vector: pExpress-1; Site: 1: EcoRV; Site 2: NotI; RNA obtained from testis tissue of 8 wk old animal. Tissues were snap-frozen and kept at -80C before RNA extraction and purification (TRI-reagent method). CDNA was primed using oligo-dT primer: 5'-pGACTAGTCTAGATCGGAGCGGCCGC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb resulted in an average insert size of 1.9 kb. This is a primary library (normalized library is NIH\_MGC\_251) and was constructed by Open Biosystems. Note: this is a NIH\_MGC library"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 2.08e-57 Length: 764  
 Score: 681.50 Matches: 140  
 Percent Similarity: 81.56% Conservative: 6  
 Best Local Similarity: 78.21% Mismatches: 21  
 Query Match: 27.92% Indels: 12  
 Gaps: 3  
 DB: 7  
 US-09-596-774-6 (1-461) x COS68334 (1-764)  
 Oy 292 ProberProvalHisPro-----ThnGlyThrSerGlnPro 303  
 Db 28 CTTCCCTCTGGGCACTCTCCAGAGGTTGCAGAAATGAATGAGCGCATCGCTCCG 87  
 Oy 304 GlnArgProGluAapCyArgProArgGlySerVallyGlyThnGlyLeu-AapPheLe 323  
 Db 88 CTGCATCTTCAAGTCAAGTCCCAAGAGC-----AAGAGCACAGAGCTTGGTCTGCT 141  
 Oy 323 uGUAAAPProlyLeuCyArgProArgGlySerVallyGlyThnGlyLeuAapPheLe 343  
 Db 142 G--GATCCCAACTGCTATATAGCTAGATGGAATCTCTTCATCTACGAGATCATGCT 198  
 Oy 343 eThAlaLeuTyLeuArgAlaTyPheSerArgSerAlaGluThrAlaAlaAsnLeuG 363

Db 199 CACGACCCCTGTACTGAGAGCAAAATTACAGAGAGTGCAGATGCTGCTTACTTCA 258  
 Oy 363 nAsPProAsnGlnLeuTyArgnGlnLeuAAsnLeuGlyAArgArgGlnGlyTyArgPvalle 383  
 Db 259 GGACCCCAACCAAGCTCTTATACGAGCTCACTTACGGCGGAGAGAGAAATATGATGTTT 318  
 Oy 383 uGUllyLeuAsnArgAlaArgAapProGluMetGlyGlyLyArgGlnGlnArgArgAapPr 403  
 Db 319 GGACAGAGAGGCGCTCGGAGATCCGAGATGGGGGCAACAGCAGAGAGAGAGAACCC 378  
 Oy 403 oGlnGlnGlyVallyTyArgnAlaLeuGlnLyAAsnLyMetAlaGlnAlaTyTserGln 423  
 Db 379 CCAGAAAGCGCTGTATCATGCTGACGAGAAAGACAAATGGCAGAGGCTTACAGTGAAT 438  
 Oy 423 eGlyThnLyArgGlyGlnArgArgArgGlyLyArgGlyVallyAsnGlyLeuTyArgGlnLyLeu 443  
 Db 439 TGGATGAAAGGCGAGAGCGGAGAGGCGAGAGGCGAGCGCTTTACCGGGGTTCTAG 498  
 Oy 443 rThAlaThrLyAsnThrTyArgAlaLeuHisMetGlnThrLeuAlaProArg 461  
 Db 499 CACTGCCACAGAGACACCTATGACGCGCTGCATATGAGACCTGCCCCCTCGC 553  
 RESULT 8  
 AGENCOURT 28621196 NIH\_MGC\_250 Rattus norvegicus CDNA clone  
 LOCUS COS72391  
 DEFINITION IMAGE:7381111 5', mRNA sequence.  
 ACCESSION COS72391  
 VERSION COS72391.1 GI:50385020  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.  
 1 (bases 1 to 786)  
 NIH-MGC http://mgc.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 Contact: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics / NIH  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgaabs-remail.nih.gov  
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical College of Wisconsin  
 CDNA Library Preparation: Open Biosystems  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLM15540 row: 1 column: 05  
 High quality sequence stop: 701.  
 Location/Qualifiers  
 1..786  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="IMAGE:7381111"  
 /lab\_host="DH10B Tona"  
 /clone\_1lb="NIH\_MGC\_250"  
 /note="Organ: thymus; Vector: pExpress-1; Site: 1: EcoRV; Site 2: NotI; RNA obtained from testis tissue of 8 wk old animal. Tissues were snap-frozen and kept at -80C before RNA extraction and purification (TRI-reagent method). CDNA was primed using oligo-dT primer: 5'-pGACTAGTCTAGATCGGAGCGGCCGC(T)25-3' and cloned into the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb resulted in an average insert size of 1.9 kb. This is a primary library (normalized library is NIH\_MGC\_251) and was constructed by Open Biosystems. Note: this is a NIH\_MGC library"  
 FEATURES  
 source

## ORIGIN

## Alignment Scores:

Pred. No.:	2,17e-57	Length:	786
Score:	681.50	Matches:	140
Percent Similarity:	81.56%	Conservative:	6
Best Local Similarity:	78.21%	Mismatches:	21
Query Match:	27.92%	Indels:	12
DB:	7	Gaps:	3

US-09-596-774-6 (1-461) x COS72391 (1-786)

```

QY 292 ProserProvalHsPro-----ThnglyThrsGlnPro 303
DB 12 CCTGCTCTGGGACACCTCCAGGGTTGCAGAAATGAAATGAGACGGATCAGTCTGC 71
QY 304 GlnArpProGlnuArpCyArpProArpGlySerVallysglyThnglyLeu-Aasphele 323
DB 72 CTGCATCTTCAAGTCAGATTCAGAGC-----AGAGCCACAGAGCTTGTGCTGCT 125
QY 323 UGluAspProLysLeuCyTyLeuLeuAspGlylleuPheileTyrglyValillel 343
DB 126 G--GATCCAAACTGCTGATATGATGATGAAATCTTCACTTACGAGATCATGCT 182
QY 343 eThralaleuTyrlleuArpAlalySpheserArpSeraIagIuthrAlaaleuLeuG1 363
DB 183 CACGCGCTGTACCTGAGAGCAAAATTCAGCAGAGTGTCAGATCTGCTTACCTTCA 242
QY 363 nAspProAsnGlnleuTyraenGluLeuAsnleuGlyArpArpGluGlyTyraSpValle 383
DB 243 GGACCCCAACAGAGCTCTATAACGAGCTCAACCTAGGGGAGAGAGAAATGATGTTT 302
QY 383 UGluLysLeuArpAlaArpAspProGluMetGlyGlyLysGlnGlnArpArpAsnPr 403
DB 303 GGAACAAGAGCGGCTCGGATCCGAGATGGGGGCAACAGACAGAGAGAGAAACC 362
QY 403 OGlnGluGlyVallyraSnalaleuGlnLysAspLysMetAlaGluAlaTyrsGlnI1 423
DB 363 CCAGAAAGGGGCTGATGATGATGATGCAAGAAAGACAGATGGCAGAGGCTTACATGAT 422
QY 423 eGlyThlysglyGlyLuarGatGatGlyLysGlyHisAspGlyLeuTyrglnGlyLeu 443
DB 423 TGGCATGAAAGGCGAGAGGCGAGAGGCGAGAGGCGACGACGCTTACACGGGTCTCAG 482
QY 443 rThnAlaThlyAspThTyraSpAlaleuHisMetGlnThrleuAlaProArp 461
DB 483 CACTGCCACCAAGACACCTATGACGCTGCAATATGACACCTGCCCTCCG 537

```

RESULT 9 993 bp mRNA linear EST 04-OCT-2001  
 LOCUS B1832527 60308128F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5221303 5',  
 DEFINITION mRNA sequence.  
 ACCESSION B1832527  
 VERSION B1832527.1 GI:15944077  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 993)  
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Contact: Robert Strausberg, Ph.D.  
 COMMENT Email: cgabs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LMNL at:  
 http://image.lnl.gov

Plate: LMNL1556 row: h column: 08  
 High quality sequence stop: 795.  
 Location/Qualifiers  
 1. 993

## FEATURES

Source  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5221303"  
 /lab\_host="DH103"  
 /clone\_id="NIH\_MGC\_120"  
 /note="Organ: pooled pancreas and spleen; Vector: pCMV-SPORT6; Site 1: HotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.:	6.78e-52	Length:	993
Score:	628.00	Matches:	125
Percent Similarity:	80.00%	Conservative:	7
Best Local Similarity:	75.76%	Mismatches:	21
Query Match:	25.73%	Indels:	12
DB:	4	Gaps:	2

US-09-596-774-6 (1-461) x B1832527 (1-993)

```

QY 296 HisProThrglyThrsGlnProGlnArpProGluAspCyArpProArpGlySerVal 315
DB 79 CATCTCGAGGACACAGTGGCCGATTACAA----- 108
QY 316 LysGlyThrglyLeuAspPheLeuGlnuAspProLysLeuCyTyrlleuLeuAspGlylle 335
DB 109 ---GGCAGAGAGCTTGCGCTGCTG---GATCCAAACTGCTTACCTGCTGATGATGATC 162
QY 336 LeuPheileTyrglyValilleThralaleuTyrlleuArpAlalySpheserArpSera 355
DB 163 CTCTTCACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 222
QY 356 AlagIuthrAlaaleuLeuGlnuAspProAsnGlnleuTyraSnalaleuLeuGlnGly 375
DB 223 GCAGAGCGCCCGGCTGACGACGAGGCGCCGACACCTCTATACAGCTCATCTACGA 282
QY 376 ArgArpGluGlyTyraSpValleuGlnLysAspAlaArpAspProGluMetGlyGly 395
DB 283 CGAAGAGAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 342
QY 396 LysGlnGlnArpArpArpArpProGlnGlnGlyVallyTyraSnalaleuGlnLysAspLys 415
DB 343 AACCCGAGAGAAAGAAACCTCTCAGAGAGGCTGTACATGAACTGCAAGAAAGATAG 402
QY 416 MetAlaGluAlaTyrsGlnI1leglyThlysglyGlyLuarGatGatGlyLysGlyHis 435
DB 403 ATGGCGAGAGGCTTACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 462
QY 436 AspGlyLeuTyrglnGlyLeuSerThralaThlyAspThTyraSpAlaleuHisMet 455
DB 463 GATGGGCTTACACGAGGCTCAGTACAGCACCAAGACACCTACAGACCCCTTACATG 522
QY 456 GlnThrleuAlaPro 460
DB 523 CAGGCGCTGCCCCC 537

```

RESULT 10 872 bp mRNA linear EST 26-MAR-2004  
 LOCUS CK629396/c 60629396F1 H03-AA0013 Apis mellifera cDNA, mRNA  
 DEFINITION AMO-AA0013-110902-011-H03 AA0013 Apis mellifera cDNA, mRNA  
 ACCESSION CK629396

VERSION CK629396.1 GI:45753871  
 KEYWORDS EST.  
 SOURCE Apis mellifera (honey bee)  
 ORGANISM Apis mellifera  
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pezomyzomata;  
 Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
 Apidae; Apis.  
 1 (bases 1 to 872)  
 REFERENCE Nunes, F.M.F., Valente, V., Sousa, J.F., Cunha, M.A.V., Pinheiro, D.G.,  
 Mata, R.M., Araujo, D.D., Costa, M.C.R., Martins, W.K., Carvalho, A.F.,  
 Mones, N., Nascimento, A.M., Peixoto, P.M.V., Silva, M.F.R.,  
 Ramos, R.G.P., Reis, L.F.L., Dias Neto, E., Souza, S.J., Simpson, A.J.G.,  
 Zago, M.A., Soares, A.E.E., Bicondi, M.M.G., Espresafico, E.M.,  
 Espindola, F.S., Peco-Larson, M.L., Simoes, Z.L.P., Hartfelder, K. and  
 Silva, W.A., Jr.  
 Open reading frame ESTs - an efficient strategy for analysis of the  
 honey bee transcriptome  
 Unpublished (2004)  
 JOURNAL Contact: Silva Jr, W. A.  
 Molecular Genetic and Bioinformatics Laboratory  
 Department of Genetics, FMRP/USP, FUNDERP  
 Rua Tenente Catão Roxo, 2501, CEP 14051-140, Brazil  
 Tel: +55 16 39639300  
 Fax: +55 16 39639309  
 Email: wilsonj@usp.br  
 This sequence was derived from the FAPESP Genome Program  
 High quality sequence start: 66  
 High quality sequence stop: 625.  
 FEATURES  
 source  
 1..872  
 /organism="Apis mellifera"  
 /mol\_type="mRNA"  
 /strain="Africanized"  
 /db\_xref="taxon:7460"  
 /sex="female, worker"  
 /dev\_stage="adult"  
 /clone\_id="AA0013"  
 /note="Organ: whole body"  
 ORIGIN  
 Alignment Scores:  
 Pred. No.: 1,12e-51 Length: 872  
 Score: 625.00 Matches: 127  
 Percent Similarity: 66.14% Conservative: 41  
 Best Local Similarity: 50.00% Mismatches: 76  
 Query Match: 25.60% Gaps: 11  
 DB: 7 Indels: 4  
 US-09-596-774-6 (1-461) x CK629396 (1-872)  
 QY 8 LeuPheLeuMetAlaAlaAlaValProLySgInIleGInLeuValGInSerGlyPro 27  
 Db 846 ATGTTCTCTTACCGCACACCGCATGCC-----AGTGGACCG 808  
 QY 28 GluLeuLysLysProGlyGluThr---ValLysIleSerCysLysAlaSer---GlyTyr 45  
 Db 807 CACCGTGGAAACGAATGTAACCTGGGCTTCATGAATTTCTCGCAAGCTTCTGCTAC 748  
 QY 46 ProPheThrAnTYrGlyMetAsnTrpValysGInAlaProGlyGInGlyLeuLysTrp 65  
 Db 747 ATCTTCACAGTTATGATATAGACTGGGTGAGCAGCCTGAACAGGAGCTTGAGAGTG 688  
 QY 66 MetGlyTrpIleAsnTrpSerThrGlyCysLysThrPheAlaAspPheLysGlyArg 85  
 Db 687 ATTGAGATGCTTTT-CCTGGAGAGGGGAGTCAATCAATGAGAGTTCAAGGGCAGG 629  
 QY 86 PheAspPheSerLeuGInTrpSerAlaAsnThrAlaTyrLeuGInIleAsnLys 105  
 Db 628 CCCACACTGAGTGGAGAGTCTCCAGCAGCTATATGAGCTCACTAGGCTGACA 569  
 QY 106 SerGluAspMetAlaThrTyrPheCysAlaArgTrpGluValTyrHISGlyTyrValPro 125  
 Db 568 TCTGAGACCTCTGCTCTATTCTGTGCTAGAGGGGACCTACTATAGCGGCTACTTGGAC 509

QY 126 TyrTrpGlyGInGlyThrThrValThrValSerSerGlyGlyGlySerGlyGly 145  
 Db 508 TTGTGGNCCAGGAGCACCGTCACCGTCTCTCCATGATGGAGNCGGTTACGGGAGGT 449  
 QY 146 GlySerGlyGlyGlyGlySerAspIleGInLeuThrGInSerHisLysPheLeuSerThr 165  
 Db 448 GGCTCTGGCGGTGGCGATCTGATCATTTAGAGTCAACCCAGTCTCCAGCAATCATGCTTCCA 389  
 QY 166 SerValIleAspArgValSerIleThrCysLysAlaSerGInAspValTyrAsnAlaVal 185  
 Db 388 TCTCCAGGGAGAGAGGTCACCATACCTGACGTGCCAGCTCAAGTATA---CGTTACATA 332  
 QY 186 AlaTrpTrpGInGInLysProGlyGInSerProLysLeuLeuIleTyrSerAlaSerSer 205  
 Db 331 TATTGTATCCACAGCAAGAGCTGATCTCTCCCGCAGACTCTCGATTATGACATCCACAC 272  
 QY 206 ArgTrpThrGlyValProSerArgPheThrGlySerGlySerGlyProAspPheThrPhe 225  
 Db 271 GTGGCTCTGGAGTCCCTTTTCGCTTCAGTGGCAGTGGGTCTGGACCTTATTCTCTC 212  
 QY 226 ThrIleSerSerValGInAlaGluAspLeuAlaValTyrPheCysGInGInHisPheArg 245  
 Db 211 ACAATCAACCGAATGAGAGCTGAGAGATGCTGCCATTATTACTGCCAGAGTGAAGTGGT 152  
 QY 246 ThrProPheThrPheGlySerGlyThrLysLeuGluLys 259  
 Db 151 TATCCGTACAGCTTCGAGAGGGGGACCAAGCTGAGAGCTGAAA 110  
 RESULT 11  
 LOCUS CV029513 618 bp mRNA linear EST 20-AUG-2004  
 DEFINITION 8329 Pull Length cDNA from the Mammalian Gene Collection Homo  
 sapiens cDNA 5' similar to BC025703, mRNA sequence.  
 ACCESSION CV029513  
 VERSION CV029513.1 GI:51487684  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 1 (bases 1 to 618)  
 REFERENCE Driscot, A., Li, N., Rosenzane-Kishikawa, T., Hao, T., Bertin, N., Li, S.,  
 Clingingsmith, T.R., Hartley, J.L., Esposito, D., Cheo, D., Moore, T.,  
 Simmons, B., Sequerra, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C.,  
 Vandenhaute, J., Quisick, M.E., Albala, J.S., Hill, D.E. and Vidal, M.  
 Human ORFeome Version 1.1: a Platform for Reverse Proteomics  
 Genome Res. (2004) In press  
 TITLE Contact: Vidal, M.  
 JOURNAL Marc Vidal Laboratory  
 Dana Farber Cancer Institute  
 1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA  
 Tel: 617 632 5180  
 Fax: 617 632 5739  
 Email: Marc.Vidal@fci.harvard.edu  
 ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF  
 template from a PCR reaction using an MGC full-length cDNA as  
 PCR primers  
 FORWARD: ATGAGTGAAGGCGCTTT  
 BACKWARD: TAGCCAGGGGCGAGGCGCTG  
 Insert Length: 618 Std Error: 29.00  
 Plate: 11027 Row: 09 Column: H  
 Seq primer: ACTGGCGTGGTATTACACAGCTGCTGACGTGGGAAAC  
 High quality sequence start: 100  
 High quality sequence stop: 617  
 POLYA=NO.  
 FEATURES  
 source  
 1..618  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"

/tissue\_type="mixed"  
/clone\_lib="Full length cDNA from the Mammalian Gene  
Collection"  
/note="Vector: mixed. The ORFs were PCR amplified from the  
MGC (Mammalian Gene Collection) as of April 2004 and  
cloned by recombinational Gateway cloning into pDONR223  
Donor Vector. Reference : MGC (Mammalian Gene Collection)  
Program Team, Generation and Initial Analysis of more than  
15,000 Full-length human and Mouse cDNA Sequences. PNAS,  
2002, 99(26), 16899-16903"

## ORIGIN

## Alignment Scores:

Pred. No.:	2,176-51	Length:	618
Score:	620.00	Matches:	126
Percent Similarity:	79.64%	Conservative:	7
Best Local Similarity:	75.45%	Mismatches:	21
Query Match:	25.40%	Indels:	13
		Gaps:	2

US-09-596-774-6 (1-461) x CV029513 (1-618)

```

QY 296 HisProthrglyThrsSerGlnProGlnArgProGlnuAspCysArgProArgGlySerVal 315
Db 30 CATCTCGAGGACAGATTGCCGATTACAGA-----59
QY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuGlySerTyrLeuLeuAspGlyI 335
Db 60 ---GGACAGAGCTTTGGCTGCTG---GATCCAAACTGCTACCTGCTGATGAT 113
QY 335 eLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSe 355
Db 114 CTTCTTATCATATGCTGTGCTATTCACCTGCTGCTGAGTGAAGTTCACAGAG 173
QY 355 ValAGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuG 375
Db 174 CGCAGACGCGCCCGCTACAGCCAGAGCCCAAGACGCTATTAACGAGCTCAATCTAGG 233
QY 375 YArgArgGluGluTyrAspValLeuGluLysArgAlaArgAspProGluMetGlyI 395
Db 234 ACGAAGAGAGAGATGATGTTTGGACAAAGACGTGGCCGGAGCCTGATGGGGG 293
QY 395 YLysGlnGlnArgArgArgAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAsp 415
Db 294 AAAGCCCGAGAGAGAGAAACCTTCAGAAAGCCTGTCATGAATGACGAGAAAGATA 353
QY 415 sMetAlaGluAlaTyrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyH 435
Db 354 GATGGCGAGGCTTACAGTGAATTTGGATGAAAGCGAGCGCGAGGGGCAAGGGGCA 413
QY 435 sAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHsme 455
Db 414 CGATGGCCTTTTACAGAGGTCTCAGTACAGCACAAGGACCTTACGAGCCTTCAAT 473
QY 455 tGlnThrLeuAlaProArg 461
Db 474 GCAAGCCTTGGCCCTCGC 492

```

RESULT 12 772 bp mRNA linear EST 04-OCT-2001  
BI838213  
LOCUS 603083131P1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5222574 5',  
DEFINITION mRNA sequence.  
ACCESSION BI838213 GI:15949763  
VERSION BI838213  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens

REFERENCE 1 (bases 1 to 772)  
AUTHORS NIH-MGC <http://mhc.nci.nih.gov/>.  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

## JOURNAL COMMENT

Unpublished (1999)  
Contact: Robert Strausberg, Ph.D.  
Email: [cgabs-remail.nih.gov](mailto:cgabs-remail.nih.gov)  
Tissue Procurement: Life Technologies, Inc.  
cDNA Library Preparation: Life Technologies, Inc.  
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
DNA Sequencing by: Incyte Genomics, Inc.  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.llnl.gov>  
Plate: LLM11559 row: m column: 07  
High quality sequence stop: 767.  
Location/Qualifiers

## FEATURES

## source

1..772  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5222574"  
/lab\_host="DH10B"  
/note="Organ: pooled pancreas and spleen; Vector:  
pCMV-SPORT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA  
source anonymous pool of spleen and pancreas from 28 yo  
male. Library is oligo-dT primed and directionally cloned  
(EcoRV site is destroyed upon cloning). Average insert  
size 1.5 kb, insert size range 1-2.5 kb. Library is  
normalized and enriched for full-length clones and was  
constructed by C. Gruber (Invitrogen). Research Genetics  
tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.:	2,996-51	Length:	772
Score:	620.00	Matches:	126
Percent Similarity:	79.64%	Conservative:	7
Best Local Similarity:	75.45%	Mismatches:	21
Query Match:	25.40%	Indels:	13
		Gaps:	2

US-09-596-774-6 (1-461) x BI838213 (1-772)

```

QY 296 HisProthrglyThrsSerGlnProGlnArgProGlnuAspCysArgProArgGlySerVal 315
Db 90 CATCTCGAGGACAGATTGCCGATTACAGA-----119
QY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuGlySerTyrLeuLeuAspGlyI 335
Db 120 ---GGACAGAGCTTTGGCTGCTG---GATCCAAACTGCTACCTGCTGATGAT 173
QY 335 eLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSe 355
Db 174 CTTCTTATCATATGCTGTGCTATTCACCTGCTGCTGAGTGAAGTTCACAGAG 233
QY 415 sMetAlaGluAlaTyrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyH 435
Db 354 GATGGCGAGGCTTACAGTGAATTTGGATGAAAGCGAGCGCGAGGGGCAAGGGGCA 413
QY 435 sAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHsme 455
Db 474 CGATGGCCTTTTACAGAGGTCTCAGTACAGCACAAGGACCTTACGAGCCTTCAAT 533

```

OY 455 tGlnThrLeuAlaProArg 461  
 DB 534 GCAGGCCCTGCCCTCGC 552

RESULT 13  
 BM921412 1038 bp mRNA linear EST 12-MAR-2002  
 LOCUS AGENCOURT 6626432 NIH\_MGC\_115 Homo sapiens CDNA clone IMAGE:5752844  
 DEFINITION 5', mRNA sequence.  
 ACCESSION BM921412  
 VERSION BM921412.1 GI:19371791  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 (bases 1 to 1038)  
 NIH-MGC http://mgi.nci.nih.gov/  
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
 JOURNAL Unpublished (1999)  
 COMMENT Contact: Robert Strusberg, Ph.D.  
 Email: cgsabbs-remail.nih.gov  
 Tissue Procurement: Life Technologies, Inc.  
 CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNLN)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNLN at:  
 http://image.lnl.gov  
 Plate: LLM12787 row: k. column: 21  
 High quality sequence stop: 730.

# FEATURES

source  
 1. 1038  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:5752844"  
 /lab\_host="DH10B"  
 /clone\_lib="NIH\_MGC\_115"  
 /note="Organ: pooled brain, lung, testis; Vector: pcwv-SPOrT6; Site\_1: NotI; Site\_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age 23-27; 1 male lung, age 27; and 1 male testis, age 69. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.8 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 021. Note: this is a NIH\_MGC library."

## ORIGIN

### Alignment Scores:

Pred. No.: 4.56e-51 Length: 1038  
 Score: 620.00 Matches: 126  
 Percent Similarity: 79.64% Conservative: 7  
 Best Local Similarity: 75.45% Mismatches: 21  
 Query Match: 25.40% Indels: 13  
 DB: 5 Gaps: 2

US-09-596-774-6 (1-461) x BM921412 (1-1038)

OY 296 HlsProthrglyThrsSerGlnProGlnArgProGlnAspGlySerAl 315  
 DB 86 CATCTCGAGGCAAGTTGGCGATTACGA----- 115  
 OY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProLeuGlySerAlaAspGly 335  
 DB 116 ---GGCACAGAGCTTTGGCTGCTG---GATCCCAAATCTGCTACCTGCGATGAAT 169  
 OY 335 eleuPhelellyrGlyVallelelThralaLeuTyrlaLeuAlaLysPheSerArgSe 355  
 DB 170 CCTCTCATCATGTGTGTCATTCACAGCTGTTCTCTGAGAGTGAAGTTACACAGAG 229

OY 355 rAlaGluThrAlaAlaasnLeuGlnAspProAsnGlnLeuTyrlaLeuAlaLeuG 375  
 DB 230 CGCAGAGCCCCCGGTACGACAGAGGCCAGAACCGCTCATATACGAGCTCAATCTAG 289  
 OY 375 YArGArGluGluTyrlaAspValleuGluLysAlaArgAlaArgAspProGluMetGly 395  
 DB 290 ACGAAGAGAGAGATGATGATGTTTGGACAGAGAGCTGGCCGGGACCTTGAGATGGGGG 349  
 OY 395 YLysGlnGlnArgArGArAspProGlnGluGlyValTyrlaAlaLeuGlnLysAsp 415  
 DB 350 AAAGCCCGAGAGAGAGAAACCTTCAGAGAGGCTGTACATGATCTGCAGAAAGATTA 409  
 OY 415 sMetAlaGluAlaTyrlaSerGluLeuGlyThrLysGlyGluArGArGlyLysGly 435  
 DB 410 GATGGCGGAGGCTTACAGTGAATTTGGGATGAAGAGGAGCGCGGAGGCGCAAGGCA 469  
 OY 435 sAspGlyLeuTyrlaGlnGlyLeuSerThrAlaThrLysAspThrTyrlaAlaLeuH 455  
 DB 470 CGATGGCTTTTACAGAGGTTCTCACTACAGCCACCAAGAGACCTTACAGCCCTTCA 529  
 OY 455 tGlnThrLeuAlaProArg 461  
 DB 530 GCAGGCCCTGCCCTCGC 548

RESULT 14  
 AL554350 1113 bp mRNA linear EST 30-MAR-2004  
 LOCUS AL554350 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens CDNA  
 DEFINITION clone CSODI082Y123 5-PRIME, mRNA sequence.  
 ACCESSION AL554350  
 VERSION AL554350.3 GI:45859108  
 KEYWORDS EST.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE  
 1 (bases 1 to 1113)  
 Li, W.B., Gruber, C., Jesssee, J. and Polayres, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished (2001)  
 COMMENT On Feb 15, 2001 this sequence version replaced gi:31276162.  
 Contact: Genoscope  
 Genoscope - Centre National de Sequencage  
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE

Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pcwvSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen. This sequence belongs to sequence cluster  
 484.x

For more information about this cluster, see  
 http://www.genoscope.cns.fr/cdna/ys=CSODI082AE120P1&c=484.r.

# FEATURES

source  
 1. 1113  
 location/Qualifiers  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CSODI082Y123"  
 /tissue type="PLACENTA COT 25-NORMALIZED"  
 /clone\_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"  
 /note="1st strand cDNA was primed with a NotI-oligo(dT)  
 primer. Five prime end enriched, double-strand cDNA was  
 digested with Not I and EcoR V sites of the pcwvSPORT 6 vector. Library was normalized."

## ORIGIN

### Alignment Scores:

Pred. No.: 5.04e-51 Length: 1113  
 Score: 620.00 Matches: 126  
 Percent Similarity: 79.64% Conservative: 7  
 Best Local Similarity: 75.45% Mismatches: 21  
 Query Match: 25.40% Indels: 13

DB: 1 Gaps: 2

US-09-596-774-6 (1-461) x AL554350 (1-1113)

QY 296 HisProthrglyThrSerGlnProGlnArgProGlnuAspCysArgProArgGlySerVal 315  
 Db 87 CATCTCGAGCGACAGTGGCCGATTACAGA-----116

QY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuGlyCysTyrLeuLeuAspGlyI 335  
 Db 117 ---GGCAGAGGCTTGGCCCTGCTG---GATCCCAAACTCTGCTACCTCGATGAGAT 170

QY 335 eleuPheleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgse 355  
 Db 171 CCTCTTCACTATAGTGTGCTCATCTCAGCTGCTTGTCTTGAGAGTGAAGTTCCAGCAGAG 230

QY 355 rAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuG 375  
 Db 231 CGCAGAGCGCCCGCGTACAGAGGCGCAGAACCGACTATACAGAGCTCAATCTAGG 290

QY 375 YArgArgGluGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyI 395  
 Db 291 ACGAAGAGAGAGTACGATGTTTGGACAGAGACGTGGCCGGAGCCCTGAGATGGGGG 350

QY 395 YLysGlnGlnArgArgAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspLys 415  
 Db 351 AAAGCCCGAGAAAGAAAGAAACCTTCAGAAAGGCTGTACATGAATGCAGAAAGATTA 410

QY 415 sMeAlaGluAlaTyrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyI 435  
 Db 411 GATGCGGAGGCTTACAGTGAATGGGATGAAAGCGAGCGGAGGCGCAAGGGGCA 470

QY 435 sAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHisMe 455  
 Db 471 CGATGGCTTTTACAGAGGTCTCAGTACGACCAAGAGACCTTACGAGCTTCACAT 530

QY 455 tGlnThrLeuAlaProArg 461  
 Db 531 GCAGGCTTGCCTCCCTGCC 549

RESULT 15  
 CR625818 1546 bp mRNA linear HTC 21-JUL-2004  
 LOCUS full-length cDNA clone CS0D005YJ16 of T cells (Turkat cell line)  
 DEFINITION Cot-10-normalized of Homo sapiens (human).  
 ACCESSION CR625818  
 VERSION CR625818.1 GI:50506625  
 KEYWORDS HTC; CNSLT cDNA.  
 SOURCE Homo sapiens (human)  
 ORGANISM Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 REFERENCE 1 (bases 1 to 1546)  
 AUTHORS Li, W.B., Gruber, C., Jessee, J. and Polayes, D.  
 TITLE Full-length cDNA libraries and normalization  
 JOURNAL Unpublished  
 REMARK Contact: Peng Liang Email: fliang@life.techn.com URL: <http://fulllength.invitrogen.com/> Invitrogen Corporation 1600 Faraday Avenue 2 (bases 1 to 1546)  
 REFERENCE Genoscope.  
 AUTHORS Direct Submission  
 TITLE Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :  
 JOURNAL BP 191 91006 EVRY cedex - FRANCE (E-mail: [secref@genoscope.cns.fr](mailto:secref@genoscope.cns.fr))  
 COMMENT - Web : [www.genoscope.cns.fr](http://www.genoscope.cns.fr)  
 1st strand cDNA was primed with a NotI-oligo(dt) primer. Five prime  
 end enriched, double-strand cDNA was digested with Not I and cloned  
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
 was normalized. Library was constructed by Life Technologies, a  
 division of Invitrogen.  
 Location/Qualifiers  
 1..1546  
 source  
 /organism="Homo sapiens"

ORIGIN  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="CS0D005YJ16"  
 /tissue\_type="T cells (Turkat cell line) Cot  
 10-normalized"  
 /plasmid="pCMVSPORT\_6"

Alignment Scores:  
 Pred. No.: 8.06e-51 Length: 1546  
 Score: 620.00 Matches: 126  
 Percent Similarity: 79.64% Conservative: 7  
 Best Local Similarity: 75.45% Mismatches: 21  
 Query Match: 25.40% Indels: 13  
 DB: 3 Gaps: 2

US-09-596-774-6 (1-461) x CR625818 (1-1546)

QY 296 HisProthrglyThrSerGlnProGlnArgProGlnuAspCysArgProArgGlySerVal 315  
 Db 69 CATCTCGAGCGACAGTGGCCGATTACAGA-----98

QY 316 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuGlyCysTyrLeuLeuAspGlyI 335  
 Db 99 ---GGCAGAGGCTTGGCCCTGCTG---GATCCCAAACTCTGCTACCTCGATGAGAT 152

QY 335 eleuPheleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgse 355  
 Db 153 CCTCTTCACTATAGTGTGCTCATCTCAGCTGCTTGTCTTGAGAGTGAAGTTCCAGCAGAG 212

QY 355 rAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuG 375  
 Db 213 CGCAGAGCGCCCGCGTACAGAGGCGCAGAACCGACTCTATACAGAGCTCAATCTAGG 272

QY 375 YArgArgGluGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyI 395  
 Db 273 ACGAAGAGAGAGTACGATGTTTGGACAGAGACGTGGCCGGAGCCCTGAGATGGGGG 332

QY 395 YLysGlnGlnArgArgAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspLys 415  
 Db 333 AAAGCCCGAGAAAGAAAGAAACCTTCAGAAAGGCTGTACATGAATGCAGAAAGATTA 392

QY 415 sMeAlaGluAlaTyrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyI 435  
 Db 393 GATGCGGAGGCTTACAGTGAATGGGATGAAAGCGAGCGCGGAGGCGCAAGGGGCA 452

QY 435 sAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHisMe 455  
 Db 453 CGATGGCTTTTACAGAGGTCTCAGTACGACCAAGAGACCTTACGAGCTTCACAT 512

QY 455 tGlnThrLeuAlaProArg 461  
 Db 513 GCAGGCTTGCCTCCCTGCC 531

Search completed: April 20, 2005, 16:33:39  
 Job time : 4534.31 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:50:16 ; Search time 65.7843 Seconds  
(without alignments)  
3588.524 Million cell updates/sec

Title: US-09-596-774-6  
Perfect score: 2441  
Sequence: 1 MAWWTLLFLPMAAKVPKQI.....LSTATKDTYDALHMOTLAPR 461

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot\_03.\*  
1: uniprot\_sprot.\*  
2: uniprot\_trembl.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	904	37.0	241	2	Q921A6 mus musculu
2	832	34.1	244	2	Q65ZC8 homo sapien
3	778	31.9	240	2	Q65ZC9 homo sapien
4	776	31.8	255	2	Q6KBD5 mus musculu
5	767	31.4	243	2	Q7TQM2 mus musculu
6	760.5	31.2	487	2	Q65ZL2 mus sp. fv/
7	748	30.6	298	2	Q9GYF0 synthetic c
8	746.5	30.6	218	2	Q925S1 mus musculu
9	725.5	29.7	248	2	Q65ZD7 mus sp. b3(
10	713	29.2	164	1	CD3Z_MOUSE
11	618	25.3	164	1	CD3Z_HUMAN
12	605	24.8	206	1	CD3H_MOUSE
13	602.5	24.7	163	1	CD3Z_PIG
14	589.5	24.1	165	1	CD3Z_RABIT
15	569	23.3	484	2	Q99LA6 mus musculu
16	539	22.1	166	1	CD3Z_SHEEP
17	498	20.4	140	2	Q65ZL3 mus sp. t91
18	486	19.9	149	1	KVSA_MOUSE
19	454.5	18.6	116	2	Q683Y7 mus musculu
20	438.5	18.0	125	2	Q6PIL0 homo sapien
21	431.5	17.7	497	2	Q8WY24 homo sapien
22	430	17.6	518	2	Q6N030 homo sapien
23	429.5	17.3	480	2	Q6P089 mus musculu
24	422.5	17.3	480	2	Q6P089 mus musculu
25	419	17.2	102	2	Q9JL79 mus musculu
26	417.5	17.1	500	2	Q6N091 homo sapien
27	417	17.1	108	2	Q8VJ10 mus musculu
28	414	17.0	500	2	Q9BRV0 homo sapien
29	413	16.9	99	2	Q9JL74 mus musculu
30	410	16.8	498	2	Q6N041 homo sapien
31	409	16.8	170	2	Q925S2 mus musculu

32	404.5	16.6	469	2	Q7Z7P5 homo sapien
33	396	16.2	470	2	Q7TMK1 mus musculu
34	395	16.2	473	2	Q9D8L4 mus musculu
35	391.5	16.0	236	2	Q6GMX8 homo sapien
36	390	16.0	136	1	KVSB_MOUSE
37	390	16.0	159	2	Q96Q50 mus musculu
38	389.5	16.0	481	2	Q91WT1 mus musculu
39	389	15.9	108	1	KVLY_HUMAN
40	388	15.9	480	2	Q8K0Z4 mus musculu
41	387.5	15.9	168	2	Q8VDC9 mus musculu
42	387.5	15.9	236	2	Q6PH7 homo sapien
43	387.5	15.9	322	2	Q6KAV0 homo sapien
44	386.5	15.8	236	2	Q6GMW1 mus musculu
45	386	15.8	613	2	Q8VCX7 mus musculu

## ALIGNMENTS

## RESULT 1

ID	Q921A6	PRELIMINARY:	PRT:	241 AA.
AC	Q921A6	01-DEC-2001 (TREMBLrel. 19, Created)		
DT	01-DEC-2001 (TREMBLrel. 19, Last sequence update)			
DT	01-MAR-2004 (TREMBLrel. 26, Last annotation update)			
DE	Anti-CEA 79 single chain Fv (fragment).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	CHUNG J.H., Choi S.J., Kim H.J., Kim I.J., Choi I.H., Lee S.D.,			
RA	Chung J.H., Choi S.J., Kim H.J., Kim I.J., Choi I.H., Lee S.D.,			
RY	Yi K.S., Suh P.G., Ryu S.H., Chung H.K.;			
RT	"Cloning and characterization of cDNAs encoding VH and VL of a			
RT	monoclonal anti-CEA antibody (CEA 79) cross-reactive with NCA-95 and			
RL	generation of a single-chain Fv molecule (scFv).";			
RL	MoJ. Cells 7:816-818(1997).			
DR	EMBL; U88067; AAB4804.1; -.			
DR	PIR; S19965; S19965.			
DR	PIR; S19967; S19967.			
DR	PIR; S19968; S19968.			
DR	HSSP; S26325; S26325.			
DR	HSSP; P01607; IBMW.			
DR	SMART; SM00406; IGV; 2.			
DR	PROSITE; PS50835; IG_LIKE; 2.			
FT	NON_TER 1			
FT	NON_TER 241			
SO	SEQUENCE 241 AA; 26086 MW; 0276887248B9C771 CRC64;			
Query Match	37.0%; Score 904; DB 2; Length 241;			
Best Local Similarity	70.2%; Pred. No. 1.7e-58;			
Matches	172; Conservative 29; Mismatches 34; Indels 10; Gaps 4;			
OY	19 QIOLVSGPELKKRGETVTKISCKASGYPTNYGMNWKQARGOGLKMKGMINTSGESTF 78			
OY	1 QVQLQQSGPELKKRGETVTKISCKASGYTFTDYGNNWVQAQGGKLGKMGWINTYGEPT 60			
OY	79 ADDPKGRFDFLETSANTAYLQINNLKSEDMATYFCARMEVYHGVVPWGGTTVTVSSG 138			
OY	61 ADPKGRFAPSLERSTASTAYLQINNLKNEEDATYFCARKDLR-YFDWGGGTTVTVSSG 119			
OY	139 GGGSGGGSGGGGGDIDLTOSHKFLSTSVGDVRSITCKASQDVYNAVAVQKQSGPK- 197			
OY	120 GGGSGGGSGGGGGDIDLTOSPSLSASLGKVTITCKASQDINKIYAWYHKDKGGRS 179			
OY	198 ---LLIYSASRRYGVSRFTGGSGGDPFTTISVQAEDLAVYFCQOHPTPTPTSGGT 254			
OY	180 AHTHTIYI---QGIPTSRSGSGGRDPSISLNEPDIATYYC-LHYDNLHTFGGT 234			
OY	255 KLEIRK 259			

Db 235 KLEIK 239

## RESULT 2

Q652C8 PRELIMINARY; PRT; 244 AA.

AC 0652C8; 25-OCT-2004 (Tremblrel. 28, Created)  
 DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
 DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
 DE Single-chain Fv (Fragment).  
 GN Name=scfv;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;  
 RX MEDLINE=97362799; PubMed=9219263;  
 RA Kontermann R.E., Wing M.G., Winter G.;  
 RT "Complement recruitment using bispecific diabodies";  
 RL Nat. Biotechnol. 15:629-631(1997).  
 DR EMBL; Y13057; CA73500.1; -.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00409; Ig; 2.  
 DR SMART; SM00406; Ig; 2.  
 DR PROSITE; PS50835; Ig\_Like; 2.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 244 AA; 26127 MW; 4B1F17868338F2BF CRC64;

Query Match 34.1%; Score 832; DB 2; Length 244;  
 Best Local Similarity 61.9%; Pred. No. 3.4e-53;  
 Matches 151; Conservative 44; Mismatches 45; Indels 4; Gaps 2;

QY 19 QIOLVOSGPELKKRGFTVKISCKASGYPTNYGMNVKQAPGQGLKRMGMINTSGESTF 78  
 DB 1 QVQLVOSGAEVKRPGDSVKVSCASGYTFSQHYHWVRQAPGQGLEWMGMDPNNGDTRF 60  
 QY 79 ADFPKGRFDESLTSANTAYLQINNLKSEDMATYFCAR--WEVYHGVPYWGQTTVTY 135  
 DB 61 AGRFQGRVMTTRDTISAAAEVWSRLNSDPTAVYYCARBEGSAIYG--MDVWGQTLVTY 119  
 QY 136 SSGGGGGGGGGGGGGSDIQLTQSHKFLSTVGVDSITCKASODVYNAVAWYQKPGQS 195  
 DB 120 SSGGGGGGGGGGGGGSDIQMTQSPSTLSASIGDRVITCRASGIGHMLAWYQKPGKA 179  
 QY 196 PKLLIYASASRYTGVPRFTSGSGSDPTFTTISVQAEIDLAVYFCQGHFRTPTFGSGTK 255  
 DB 180 PKLLIYASASLASGAPRFSGSGSDPTFTTISLQPDPAIYTCQGYSNVPLTFGGGK 239  
 QY 256 LEIK 259  
 DB 240 LEIK 243

## RESULT 3

ID 0652C9 PRELIMINARY; PRT; 240 AA.

AC 0652C9; 25-OCT-2004 (Tremblrel. 28, Created)  
 DT 25-OCT-2004 (Tremblrel. 28, Last sequence update)  
 DT 25-OCT-2004 (Tremblrel. 28, Last annotation update)  
 DE Single-chain Fv (Fragment).  
 GN Name=scfv;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
 NCBI\_TaxID=9606;

RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C1q/7;  
 RX MEDLINE=97362799; PubMed=9219263;  
 RA Kontermann R.E., Wing M.G., Winter G.;  
 RT "Complement recruitment using bispecific diabodies";  
 RL Nat. Biotechnol. 15:629-631(1997).  
 DR EMBL; Y13056; CA73499.1; -.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig\_v.  
 DR Pfam; PF00047; Ig; 2.  
 DR SMART; SM00409; Ig; 2.  
 DR PROSITE; PS50835; Ig\_Like; 2.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 240 AA; 25569 MW; FDCFD3645F64B373 CRC64;

Query Match 31.9%; Score 778; DB 2; Length 240;  
 Best Local Similarity 60.3%; Pred. No. 3.1e-49;  
 Matches 146; Conservative 39; Mismatches 53; Indels 4; Gaps 2;

QY 19 QIOLVOSGPELKKRGFTVKISCKASGYPTNYGMNVKQAPGQGLKRMGMINTSGESTF 78  
 DB 1 QVQLVOSGGGLVPGGSLRLSCAASGFTTISYGMHWVRQAPGQGLMVAIVSDGSNKYY 60  
 QY 79 ADFPKGRFDESLTSANTAYLQINNLKSEDMATYFCAR--WEVYHGVPYWGQTTVTY 137  
 DB 61 ADFVKGRTTISRDNSTKNTLYLQWNSLRADPTAVYYCARQW--GDSLDPMGKQTLVTY 117  
 QY 138 GGGGGGGGGGGGGSDIQLTQSHKFLSTVGVDSITCKASODVYNAVAWYQKPGQSPK 197  
 DB 118 GGGGGGGGGGGGGSDIQMTQSPSTLSASIGDVTITCRASGIGHMLAWYQKPGKAPK 177  
 QY 198 LLTIYASASRYTGVPRFTSGSGSDPTFTTISVQAEIDLAVYFCQGHFRTPTFGSGTKLE 257  
 DB 178 LLTIYASASLASRAPRFSGSGSDPTFTTISLQPDPAIYTCQGYSNVPLTFGGGKLE 237  
 QY 258 IK 259  
 DB 238 IK 239

## RESULT 4

ID 06KB05 PRELIMINARY; PRT; 255 AA.

AC 06KB05; 05-JUL-2004 (Tremblrel. 27, Created)  
 DT 05-JUL-2004 (Tremblrel. 27, Last sequence update)  
 DT 05-JUL-2004 (Tremblrel. 27, Last annotation update)  
 DE scfv B8E5 protein (Fragment).  
 GN Name=scfv B8E5;  
 OS Mus musculus (mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Balb/c;  
 RA Briand J.P., Hoebeke J.;  
 RL Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.  
 DR EMBL; AJ746180; CAG34081.1; -.  
 DR HSP; P01837; 1KCR.  
 DR InterPro; IPR003599; Ig.  
 DR InterPro; IPR007110; Ig-like.  
 DR InterPro; IPR003596; Ig\_v.  
 DR SMART; SM00409; Ig; 2.  
 DR SMART; SM00406; Ig; 2.  
 DR PROSITE; PS50835; Ig\_Like; 2.  
 FT NON\_TER 1  
 FT NON\_TER 1  
 SQ SEQUENCE 255 AA; 27445 MW; B68BD38395DP713B CRC64;

Query Match 31.8%; Score 776; DB 2; Length 255;  
 Best Local Similarity 59.8%; Pred. No. 4.8e-49;  
 Matches 149; Conservative 36; Mismatches 56; Indels 8; Gaps 2;

QY 19 QIOLVOSGPELKKPEYTVKISCKASGYPTNTYGMNVAQAPQGLKMMGMINSTGESETF 78  
 DB 1 QVQLQQSGGDLVKKPGASLKVSCAAGSFTSSYGMSWVRGTPDKRLAEVATITSGSYTY 60  
 QY 79 ADPFGRDPFSLSTANTAYLIQINNLSKSDMATYCARV-EVYGYVPYWGCTTVVS 136  
 DB 61 PDSVGRFTTISRDNKNTLYLQMSLSKSEDTAMYYCARHINRYDGAEDYWGCTTVVS 120  
 QY 137 SGGSGGGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDYVNA-----VAMTQQ 190  
 DB 121 SGGSGGGGGGGGGSDIYMAQSPSSLSVAGAEKIMCKSKSQSLNSGNQNYLAWTQQ 180  
 QY 191 KPGQSPKLLIYASASRYTGVPSPRFTGSGGPDFTTISVQAEDLAVYFCQGHFTPTPTF 250  
 DB 181 KPGQSPKLLIYASASRYTGVPSPRFTGSGGPDFTTISVQAEDLAVYFCQGHFTPTPTF 240  
 QY 251 GSGTLEIK 259  
 DB 241 GAGTLEIK 249

RESULT 5

ID Q7QM2 PRELIMINARY; PRT; 243 AA.  
 AC Q7QM2;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Scfv 6H8 protein (Fragment).  
 GN Name=bcfv 6H8;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Balb/C;  
 RA MEDLINE=22853226; PubMed=12860977; DOI=10.1074/jbc.M306877200;  
 RT Peter J.C., Fekethari P., Billiard P., Wallukat G., Heebke J.;  
 RT "scfv single chain antibody variable fragment as inverse agonist for  
 the beta-2 adrenergic receptor";  
 RL J. Biol. Chem. 278:36740-36747(2003).  
 DR EMBL; AJ574851; CAE00495.1; -.  
 DR HSBP; P01751; IAGM.  
 DR InterPro: IPR007110; IG-like.  
 DR InterPro: IPR003596; IG\_v.  
 DR SMART; SM00406; IGV; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 2.  
 FT NON TER 1  
 SQ SEQUENCE 243 AA; 25976 MW; BFFPF64D2DCAF76 CRC64;

Query Match 31.4%; Score 767; DB 2; Length 243;  
 Best Local Similarity 60.6%; Pred. No. 2.1e-48;  
 Matches 146; Conservative 42; Mismatches 49; Indels 4; Gaps 3;

QY 19 QIOLVOSGPELKKPEYTVKISCKASGYPTNTYGMNVAQAPQGLKMMGMINSTGESETF 78  
 DB 1 QVQLQQSGGDLVKKPGASLKVSCAAGSFTSSYGMSWVRGTPDKRLAEVATITSGSYTY 60  
 QY 79 ADPFGRDPFSLSTANTAYLIQINNLSKSDMATYCARV-EVYGYVPYWGCTTVVS 138  
 DB 61 DEKFNKGLTVDTSSSTAYMHLSSLASDSAVYCARGG--RG-LDVMGAGTTLTVSSG 117  
 QY 139 GGGSGGGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDYVNAVAMTQQPQSKRL 198  
 DB 118 GGGSGGGGGGGGGSDIQMTQSSSSPSVSLGDRVITTCASBDIYNRLAMTQPKPNARL 177  
 QY 199 LYSASRYTGVPSPRFTGSGGPDFTTISVQAEDLAVYFCQGHFTPTPTFGSGTLEIK 258

DB 178 LISGATSLSTGVSPRFSGSGKDYTLSTLQEDVATVYCCQYWSYR-TFGGCTLEIK 236  
 QY 259 K 259  
 DB 237 K 237

RESULT 6

ID Q65ZL2 PRELIMINARY; PRT; 487 AA.  
 AC Q65ZL2;  
 DT 25-OCT-2004 (TREMBLrel. 28, Created)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last sequence update)  
 DT 25-OCT-2004 (TREMBLrel. 28, Last annotation update)  
 DE FV/M4.  
 GN Name=M4-IFN-ctau>;  
 OS Mus sp.  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10095;  
 RX [1]  
 RP SEQUENCE FROM N.A.  
 RC MEDLINE=96272580; PubMed=8688499;  
 RA Qi Y., Xiang J.;  
 RT "A genetically engineered single-gene-encoded anti-TAG72 chimeric  
 antibody secreted from myeloma cells";  
 RL Hum. Antibodies Hybridomas 6:161-166(1995).  
 DR EMBL; S82493; AAB37424.2; -.  
 DR InterPro: IPR003599; IG.  
 DR InterPro: IPR007110; IG-like.  
 DR InterPro: IPR003597; IG\_c1.  
 DR InterPro: IPR003598; IG\_c2.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR InterPro: IPR003596; IG\_v.  
 DR Pfam; PF07654; CI-sec; 2.  
 DR SMART; SM00407; IG; 4.  
 DR SMART; SM00409; IG; 3.  
 DR SMART; SM00407; IGC1; 2.  
 DR SMART; SM00408; IGC2; 2.  
 DR SMART; SM00406; IGV; 2.  
 DR PROSITE; PS50835; IG\_LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN 1.  
 SQ SEQUENCE 487 AA; 53578 MW; CTBA69F3055504 CRC64;

Query Match 31.2%; Score 760.5; DB 2; Length 487;  
 Best Local Similarity 47.8%; Pred. No. 1.5e-47;  
 Matches 150; Conservative 55; Mismatches 78; Indels 31; Gaps 6;

QY 1 MAMVWTLFLMAAK-VPKQIOLVOSGPELKKPEYTVKISCKASGYPTNTYGMNVAQAP 59  
 DB 1 MEMSWFLFLSLVTTGHSQVQLQSSDAELVKKPGASLKVSCKASGYTTTDAIIMAKQK 60  
 QY 60 GQGLKMMGMINSTGESETFADPFGRDPFSLSTANTAYLIQINNLSKSDMATYCARWEV 119  
 DB 61 EQLLEMGYISPGNDIKYNEKFKGKATLTADKSSSTAYMGLNSLSDSAVYFCKR--- 117  
 QY 120 YHGVVPYWGCTTVVS SGGGGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQ 179  
 DB 118 --SYGHWGCTTTLT-SGGGGGGGGGGSGGSGSRIOQTQSPASLSVSGELVITTCRASE 174  
 QY 180 DVNAVAMTQKPPQSPKLLIYASASRYTGVPSPRFTGSGGPDFTTISVQAEDLAVYFC 239  
 DB 175 NYSNLAMTQKQKQSPQVLVATNLADGVSPRFSGSGSTQYSLKINSLSQSEDFSGY 234  
 QY 240 CQHFRTPTFGSGTLEIKALISNSVMYFSSVPLQKYNSTTKRVLTPSPV---- 295  
 DB 235 CQHFMGTPPTFGSGTLEIK-----PSKSCDKHTGCP---PQAPBELT 275  
 QY 296 -HPTGTSOPARPED 308  
 DB 276 GGPVFLFPKPKD 289

```

RESULT 7
ID 090YFO PRELIMINARY; PRT; 298 AA.
AC 090YFO:
DT 01-MAY-2000 (TREMblrel. 13, Created)
DT 01-MAY-2000 (TREMblrel. 13, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE CN 8 single chain antibody.
DE Name=CN 8 scFv;
OS synthetic construct.
OC other sequences; artificial sequences.
OX NCBI_TaxID=32630;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=20183931; PubMed=10706631; DOI=10.1073/pnae.050582197;
RA Shinohara N., Demura T., Fukuda H.;
RT "Isolation of a vascular cell wall-specific monoclonal antibody
RT recognizing a cell polarity by using a phase display subtraction
RT method.";
RL Proc. Natl. Acad. Sci. U.S.A. 97:2585-2590(2000).
DR EMBL; AB036341; BAAB8633.1; -.
DR PIR; A33933; A33933.
DR PIR; S19112; S19112.
DR HSSP; P01820; 1A70.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PSS0835; IG LIKE; 2.
SQ SEQUENCE 298 AA; 31867 MW; E0F96B8A17004317 CRC64;

Query Match 30.6%; Score 748; DB 2; Length 298;
Best Local Similarity 52.5%; Pred. No. 6.7e-47;
Matches 145; Conservative 46; Mismatches 77; Indels 8; Gaps 2;

OY 7 LLFLM-----AAAKVKQIOLVSGPBLKKPBTYVISCASGYPTTYGMMVYKQAP 59
DB 21 LLFLIPLVPPYPAQPMAYKLOQSGGLVKGSLSCAASGSDPSRYSWMSVYQAP 80
OY 60 GQGLKMGWINTSGESTFADDFKGRFDFSLTSANTAYLOINNKKSEDMATYFCARWEV 119
DB 81 GKGLWEMIGEINPDSSTINYPSTLKDKFLISRDNKNTLYLQMSKVRSEDTALYICAR-AS 139
OY 120 YHGVVPWGGQTTVTVSSGGGGGGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQ 179
DB 140 YHGSAAWGGQTTVTVSSGGGGGGGGGGSDIEILTQSPASISASVGEVITTCRASG 199
OY 180 DVYNAVAMWYQKPGQSPKLLIYSASRYTGVPSPFTSGSGPDTFTIISVQAE DLAVF 239
DB 200 NIHYILAMWYQKQKQKSPQLLVYNAKTLADGVPSKFSGSGSTQYSLKINSLOPEDFGSY 259
OY 240 CQHFRTPTFTFGSGTKLEIKALEISNVYFSSVVP 275
DB 260 CQHFMTPTFTFGSGTKLEIKRAAGAPVPPDPLEP 295

RESULT 8
ID 0925S1 PRELIMINARY; PRT; 218 AA.
AC 0925S1:
DT 01-DEC-2001 (TREMblrel. 19, Created)
DT 01-DEC-2001 (TREMblrel. 19, Last sequence update)
DT 01-OCT-2003 (TREMblrel. 25, Last annotation update)
DE MKPS (Fragment).
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN (1)
RP SEQUENCE FROM N.A.
RX STRAIN=BALB/c;
RX PubMed=11819679;
RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,

```

```

RA Su C.;
RT "Mechanism of exogenous nucleic acids and their precursors improving
RT the repair of intestinal epithelium after irradiation in mice.";
RL World J. Gastroenterol. 6:709-717(2000).
RN (2)
RP SEQUENCE FROM N.A.
RC STRAIN=BALB/c;
RA Cui D., Zeng G., Yan X., Li X., Su C.;
RT "Cloning of mouse genes related to repairing of intestinal epithelium
RT of the irradiated mice by treatment with the intestinal RNA of mice of
RT the same strain.";
RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).
DR EMBL; AF240168; AAK43733.1; -.
DR HSSP; P01665; 1ONZ.
DR SMART; SM00406; IGV; 1.
DR PROSITE; PSS0835; IG LIKE; 1.
FT NON TER 218
SQ SEQUENCE 218 AA; 23013 MW; 527E4FA8F7982817 CRC64;

Query Match 30.6%; Score 746.5; DB 2; Length 218;
Best Local Similarity 66.2%; Pred. No. 5.8e-47;
Matches 141; Conservative 28; Mismatches 39; Indels 5; Gaps 2;

OY 19 QIOLVSGPBLKKPBTYVISCASGYPTTYGMMVYKQAPGQGLKMGWINTSGESTP 78
DB 3 QVXLQSGPBLKKPBTYVISCASGYTTTACQWQKMPGGLKMWIGWINTSGVPKY 62
OY 79 ADPFKGRFDFSLTSANTAYLOINNKKSEDMATYFCARWEVHYGYPVYGGQTTVSSG 138
DB 63 ABEFKGRFAPSLTSLSTALQISLNKNEDTATYFCMRND-YDGGFAYWGGQTTVSSG 121
OY 139 GGGSGGGSGGGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNA-----VAVYQKPGQ 194
DB 122 GGGSGGGSGGGGGSDIQLTQSPASISASVGEVITTCRASGSVDNIGISFNMWFOQKPGQ 181
OY 195 SPKLLIYSASRYTGVPSPFTSGSGPDTFTIISVQAE DLAVF 227
DB 182 PPKLLIYAAKSGSGVPAGLASGSGTDFSLNI 214

RESULT 9
ID 065207 PRELIMINARY; PRT; 248 AA.
AC 065207:
DT 25-OCT-2004 (TREMblrel. 28, Created)
DT 25-OCT-2004 (TREMblrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMblrel. 28, Last annotation update)
DE B3 (Fv)-PB40 (Fragment).
GN Name=B3 (Fv)-PB40;
OS Mus sp.
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10095;
RN (1)
RP SEQUENCE FROM N.A.
RX MEDLINE=92020904; PubMed=1924323;
RX Brinkmann U., Pal L.H., Fitzgerald D.J., Williamson M., Pastan I.;
RT "B3 (Fv)-PB38KDEL, a single-chain immunotoxin that causes complete
RT regression of a human carcinoma in mice.";
RL Proc. Natl. Acad. Sci. U.S.A. 88:8616-8620(1991).
DR EMBL; S57990; AAB1971.2; -.
DR InterPro; IPR003599; Ig.
DR InterPro; IPR007110; Ig-like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; IG; 2.
DR SMART; SM00406; IGV; 2.
DR PROSITE; PSS0835; IG LIKE; 2.
FT NON TER 248
SQ SEQUENCE 248 AA; 26634 MW; 7A3759B43B570950 CRC64;

Query Match 29.7%; Score 725.5; DB 2; Length 248;
Best Local Similarity 55.3%; Pred. No. 2.4e-45;

```

Matches	136;	Conservative	42;	Mismatches	63;	Indels	5;	Gaps	1
Qy	20	IQLVQSGPELKKPGTETVXISCKASGYPTPTNGMNVKQAPGGGLKMGMINSTSGSTFA	79						
Db	3	VKLVESSGGGLVQPGSGSLKSLSCATSGFTFDYVMYWRQPEKRLBMAVYISNDSSAAYS	62						
Qy	80	DDFKRFPSPLETSANTAYLQINNLKSEMMATYFCARMEVHGVYPYMGQITTYVSSGG	139						
Db	63	DTVGRFRFISIDNANNTLYLQWRSRKISEPTAYISCARGLAMGAPAYMGQGLTYVSSGG	122						
Qy	140	GGSGGGGGGGGGSDIQLTQSHKFEIYSTVGDRISINCKASQDVYVN-----VAMVQOKPQ	194						
Db	123	GGSGGGGGGGGGSDVLMQTPSLSPSLVSLADQASISCRSQITIVHNGNTYLMWYLOKPEQ	182						
Qy	195	SPKLTIYASSRNYTGVPSRFRFTSGSGSPDFTPTTIVSSVOAEDLAVYECOOHFRPTFTGSGT	254						
Db	183	SPKLTIYKVSNRFSGVPPDRFSSGSGTDPTLTLSVVAEDLGVYICFGSHNPFPGSGT	242						
Qy	255	KLEIYA 260							
Db	243	KLEIYA 248							
RESULT 10									
CD32_MOUSE		STANDARD:	PRT:	164	AA.				
ID	CD32_MOUSE	STANDARD:	PRT:	164	AA.				
AC	P24161; Q9D3G3;								
DT	01-MAR-1992 (Rel. 21, Created)								
DT	01-MAR-1992 (Rel. 21, Last sequence update)								
DT	25-OCT-2004 (Rel. 45, Last annotation update)								
DE	T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor								
DE	T3 zeta chain).								
OS	Name=CD3z; Synonyms=TCrz;								
OS	Mus musculus (Mouse).								
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;								
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.								
OX	NCBI_TaxID=10090;								
RN	[1]								
RP	SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.								
RX	MEDLINE=86145643; PubMed=3278377;								
RA	Weisman A.M., Banlyash M., Hou D., Samelson L.E., Burgess W.H.,								
RA	Klausner R.D.;								
RT	"Molecular cloning of the zeta chain of the T cell antigen receptor.,"								
RL	Science 239:1018-1021(1988).								
RN	[2]								
RP	SEQUENCE FROM N.A.								
RC	TISSUE=Liver;								
RC	MEDLINE=89327299; PubMed=2787796;								
RA	Banlyash M., Hsu V.W., Seldin M.F., Klausner R.D.;								
RT	"The isolation and characterization of the murine T cell antigen								
RT	receptor zeta chain gene.,"								
RL	J. Biol. Chem. 264:13252-13257(1989).								
RN	[3]								
RP	SEQUENCE FROM N.A.								
RC	STRAIN=C57BL/6J; TISSUE=Thymus;								
RC	MEDLINE=22354683; PubMed=12468851; DOI=10.1038/nature01266;								
RA	Okazaki Y., Furuno M., Kasukawa T., Adachi Y., Bono H., Kondo S.,								
RA	Nikaido I., Otsu N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,								
RA	Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,								
RA	Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.W.,								
RA	Schirral L.M., Kanapin A., Matveia H., Batalov S., Betzel K.W.,								
RA	Blake J.A., Bradt D., Bruscia V., Chochia C., Corbani L.E., Cousins S.,								
RA	Dalla E., Dargatz T.A., Fletcher C.F., Forrest A., Frizer K.S.,								
RA	Garmestland T., Gariboldi M., Giersi C., Godzik A., Gough J.,								
RA	Grimmond S., Guerinich S., Hirokawa N., Jackson I.J., Jarvis E.D.,								
RA	Kanai A., Kawai H., Kawasawa Y., Kedzierski R.W., King B.L.,								
RA	Kanagaya A., Kurochkin I.V., Lee Y., Lennard B., Lyons P.A.,								
RA	Majloot D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,								
RA	Nakashima T., Numata K., Okido T., Pavan W.J., Pereira G., Pesole G.,								
RA	Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,								
RA	Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,								
RA	Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,								
RA	Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,								

RA Vercario R., Wagner L., Wahlstedt C., Wang Y., Marande Y., Wells C.,  
RA Wilming L.G., Wyman-Boris A., Yanagisawa M., Yang I., Yang L.,  
RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carinini P., Hayatsu N.,  
RA Hirozane-Kishikawa T., Konno H., Nakamura M., Sakazume N., Sato K.,  
RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
RA Hara A., Hashimoto W., Imocani K., Jibit Y., Itoh M., Kagawa I.,  
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
RA Yamunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
RA Birney E., Hayashizaki Y.;  
RT "Analysis of the mouse transcriptome based on functional annotation of  
RL 60,770 full-length cDNAs.";  
RL Nature 420:563-573(2002).  
RN [4]  
RN SEQUENCE FROM N.A.  
RC STRAIN=C57BL/6; TISSUE=Hematopoietic;  
RC MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RX Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helt P.,  
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Ussid T.B., Toshiyuki S., Carinini P., Prange C.,  
RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,  
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
RA Richard S.K., Morley R.C., Hale S., Garcia A.W., Gay L.J., Hulyk S.W.,  
RA Villalón D., Wozny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Fahey J., Heltan E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
RA Buttefield Y.S.N., Krzywinski M.T., Skalske U., Smalins D.E.,  
RA Schenck A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RL and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
RN [5]  
RN PARTIAL SEQUENCE.  
RX MEDLINE=90239005; PubMed=2139725;  
RX Jin Y.J., Clayton L.K., Howard F.D., Koyasu S., Stieh M.,  
RA Steibrich R., Tarr G.E., Reinherz E.L.;  
RT "Molecular cloning of the CD3 epsilon subunit identifies a CD3 zeta-  
RT related product in thymus-derived cells.";  
RL Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323(1990).  
RN [6]  
RN ALTERNATIVE SPLICING.  
RX PubMed=2150596.  
RX Ohno H., Saito T.;  
RA "CD3 zeta and eta chains are produced by alternative splicing from a  
RT common gene.";  
RL Int. Immunol. 2:1117-1119(1990).  
RN [7]  
RN ERRATUM.  
RX Ohno H., Saito T.;  
RA Int. Immunol. 4:1339-1339(1992).  
RN [8]  
RN INTERACTION WITH SLA.  
RX MEDLINE=20130290; PubMed=10662792;  
RX Sosinowski T., Pandey A., Dixit V.M., Weiss A.;  
RA "Src-like adaptor protein (SLAP) is a negative regulator of T cell  
RT receptor signaling.";  
RL J. Exp. Med. 191:463-474(2000).  
RN [9]  
RN INTERACTION WITH SLA2.  
RX MEDLINE=22013997; PubMed=11891219; DOI=10.1074/jbc.M10318200;  
RX Pandey A., Ibarrola N., Krachmarova I., Fernandez M.M.,  
RA Constantinescu S.N., Ohara O., Sawaadikosol S., Lodish H.F., Mann M.;  
RT "A novel Src homology 2 domain-containing molecule, Src-like adaptor  
RT protein-2 (SLAP-2), which negatively regulates T cell receptor  
RT signaling.";  
RL J. Biol. Chem. 277:19131-19136(2002).  
CC -!- FUNCTION: Probable role in assembly and expression of the TCR  
CC complex as well as signal transduction upon antigen triggering.

CC -1- SUBUNIT: Interacts with DOKK2 (By similarity). The TCR/CD3 complex  
 CC of T lymphocytes consists of either a TCR alpha/beta or TCR  
 CC gamma/delta heterodimer coexpressed at the cell surface with the  
 CC invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and  
 CC eta. CD3-zeta forms either homodimers or heterodimers with CD3-  
 CC eta. Interacts with SLA and SLA2.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=CD-3-zeta;  
 CC IsoId=P24161-1; Sequence=Displayed;  
 CC Name=CD-3-eta;  
 CC IsoId=P29020-1; Sequence=External;  
 CC -1- PFM: Phosphorylated on Tyr residues after T-cell receptor  
 CC triggering.  
 CC -1- SIMILARITY: Belongs to the CD3z/FCER1G family.  
 CC -1- SIMILARITY: Contains 3 ITAM domains.  
 CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 CC EMBL: M19729; AAA0171.1; -;  
 CC EMBL: J04967; AAA50301.1; -;  
 CC EMBL: AK017904; BAB30997.1; -;  
 CC EMBL: BC052824; AAH52824.1; -;  
 CC PIR: A40104; A40104.  
 CC MGI: MGI:88334; C33z.  
 CC InterPro: IPR003110; ITAM.  
 CC Pfam: PF02189; ITAM; 3.  
 CC Alternative splicing: Direct protein sequencing; Phosphorylation;  
 CC Receptor; Repeat; Signal; T-cell; Transmembrane.  
 CC SIGNAL 1 21  
 CC CHAIN 22 164 T-cell surface glycoprotein CD3 zeta  
 CC DOMAIN 22 30 chain.  
 CC TRANSMEM 31 51 Extracellular (Potential).  
 CC DOMAIN 52 164 Potential.  
 CC DOMAIN 69 89 Cytoplasmic (Potential).  
 CC DOMAIN 108 129 ITAM 1.  
 CC DOMAIN 139 159 ITAM 2.  
 CC DISULFID 32 32 ITAM 3.  
 CC MOD RES 153 153 Interchain (Potential).  
 CC CONFLICT 153 153 Phosphotyrosine.  
 CC SEQUENCE 164 AA; 18637 MW; 188022035A312831 CRC64;

Query Match 29.2%; Score 713; DB 1; Length 164;  
 Best Local Similarity 99.3%; Pred. No. 1.2e-44;  
 Matches 138; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 323 LEDPKLCYLDGLIFITGVITITALLYLPAKFSRSATANTADPNQLYNELNLRREBYDV 382  
 DB 26 LLDPRCLCYLDGLIFITGVITITALLYLPAKFSRSATANTADPNQLYNELNLRREBYDV 85  
 QY 383 LEKRRADPEWKGKQRRNPQESGVNALOKDKAEVYSITGKRRRKKHGDGLYOGI 442  
 DB 86 LEKRRADPEWKGKQRRNPQESGVNALOKDKAEVYSITGKRRRKKHGDGLYOGI 145  
 QY 443 STATKDTYDALHMQTLAPR 461  
 DB 146 STATKDTYDALHMQTLAPR 164

RESULT 11  
 ID CD3Z\_HUMAN STANDARD; PRT; 164 AA.  
 AC P20963; O8TAX4;  
 DT 01-FEB-1991 (Rel. 17, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)

DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor  
 DE T3 zeta chain).  
 GN Name=CD3z; Synonyms=73z, TCRz;  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OK NCBI\_TaxId=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RX MEDLINE=89071765; PubMed=2974162;  
 RA Weissman A.M., Hou D., Orloff D.G., Modi W.S., Suarez H.,  
 RA O'Brien S.J., Klausner R.D.,  
 RT "Molecular cloning and chromosomal localization of the human T-cell  
 RT receptor zeta chain: distinction from the molecular CD3 complex";  
 RL Proc. Natl. Acad. Sci. U.S.A. 85:9705-9713 (1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RX TISSUE=Pancreas;  
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Bueltow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueudin T.B., Toshlyuk S., Carninci P., Prange C.,  
 RA Rata S.S., Loguettano N.A., Peters J., Abramson R.D., Mullaly S.J.,  
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hultk S.W.,  
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green S.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield J.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,  
 RA Schnerich A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [3]  
 RP INTERACTION WITH HIV-1 NFP.  
 RX PubMed=10224289;  
 RA Xu X.-N., Laferriere B., Screaton G.R., Kraft M., Wolf D., Kolanus W.,  
 RA Mongkolkeha J., McMichael A.J., Baur A.S.;  
 RT "Induction of Fas ligand expression by HIV involves the interaction of  
 RT Nef with the T cell receptor zeta chain";  
 RL J. Exp. Med. 189:1489-1496 (1999).  
 RN [4]  
 RP INTERACTION WITH SLA.  
 RX MEDLINE=99380595; PubMed=10449770; DOI=10.1073/pnas.96.17.9775;  
 RA Tang J., Sawadikowski S., Chang J.-H., Burckoff S.J.;  
 RT "SLAP, a dimeric adapter protein, plays a functional role in T cell  
 RT receptor signaling";  
 RL Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780 (1999).  
 RN [5]  
 RP INTERACTION WITH DOKK2.  
 RX MEDLINE=22165501; PubMed=12176041; DOI=10.1016/S0006-291X(02)00931-2;  
 RA Nishihara H., Maeda M., Tsuda M., Makino Y., Sawa H., Nagashima K.,  
 RA Tanaka S.;  
 RT "DOKK2 mediates T cell receptor-induced activation of Rac2 and IL-2  
 RT transcription";  
 RL Biochem. Biophys. Res. Commun. 296:716-720 (2002).  
 RN [6]  
 RP PHOSPHORYLATION SITES TYR-83; TYR-111; TYR-123 AND TYR-142.  
 RX PubMed=12522270; DOI=10.1073/pnas.2436191100;  
 RA Salomon A.R., Ficarro S.B., Brill L.M., Brinker A., Phung Q.T.,  
 RA Ericson C., Sauer K., Brock A., Horn D.M., Schultz P.G., Peters E.C.;  
 RT "Profiling of tyrosine phosphorylation pathways in human cells using  
 RT mass spectrometry";  
 RL Proc. Natl. Acad. Sci. U.S.A. 100:443-448 (2003).  
 RN [7]  
 RP STRUCTURE BY NMR OF 136-149.

RX MEDLINE=93201600; PubMed=7680960; DOI=10.1016/0092-8674(93)90405-F;  
 RA Wakeman G., Shoelson S.E., Pant N., Cowburn D., Kurlyan J.;  
 RT "Binding of a high affinity phosphotyrosyl peptide to the Src SH2  
 RT domain: crystal structures of the complexed and peptide-free forms.";  
 RL Cell 72:779-790(1993).  
 CC -1- FUNCTION: Probable role in assembly and expression of the TCR  
 CC complex as well as signal transduction upon antigen triggering.  
 CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a  
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the  
 CC cell surface with the invariant subunits of CD3 labeled gamma,  
 CC delta, epsilon, zeta, and eta. CD3-zeta forms either homodimers or  
 CC heterodimers with CD3-eta. Interacts with SLA and SLA2. Interacts  
 CC with DOCK2. Interacts with HIV-1 Nef protein.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=CD-3-zeta;  
 CC IsoId=P20963-1; Sequence=Displayed;  
 CC Name=CD-3-eta;  
 CC IsoId=P20963-2; Sequence=Not described;  
 CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor  
 CC triggering (by similarity).  
 CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.  
 CC -1- SIMILARITY: Contains 3 ITAM domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL collaboration -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).  
 CC -----  
 DR EMBL; J04132; AAA60394.1; -;  
 DR EMBL; BC025703; AAH25703.1; -;  
 DR PIR; A31768; A31768.  
 DR PDB; 1TCE; NMR; B=136-149.  
 DR Genew; HGNC:1677; CD3Z.  
 DR H-InVDB; HIX001296; -;  
 DR MIM; 186780; -;  
 DR GO; GO:0005886; C:plasma membrane; TAS.  
 DR GO; GO:0042101; C:T-cell receptor complex; TAS.  
 DR GO; GO:0005515; F:protein binding; IPI.  
 DR GO; GO:0042803; F:protein homodimerization activity; NAS.  
 DR InterPro: IPR003110; ITAM.  
 DR Pfam; PF02189; ITAM; 3.  
 DR SMART; SM00077; ITAM; 3.  
 KW 3D-structure; Alternative splicing; Phosphorylation; Receptor; Repeat;  
 KW Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 21  
 FT CHAIN 22 164 T-cell surface glycoprotein CD3 zeta  
 FT chain.  
 FT DOMAIN 22 30 Extracellular (Potential).  
 FT TRANSMEM 31 51 Potential.  
 FT DOMAIN 52 164 Cytoplasmic (Potential).  
 FT DOMAIN 69 89 ITAM 1.  
 FT DOMAIN 108 129 ITAM 2.  
 FT DOMAIN 139 159 ITAM 3.  
 FT DISULFID 32 32 Interchain (Potential).  
 FT MOD\_RES 83 83 Phosphotyrosine.  
 FT MOD\_RES 111 111 Phosphotyrosine.  
 FT MOD\_RES 123 123 Phosphotyrosine.  
 FT MOD\_RES 142 142 Phosphotyrosine.  
 FT MOD\_RES 153 153 Phosphotyrosine (by similarity).  
 FT CONFLICT 60 61 DA -> EP (in Ref. 1).  
 FT CONFLICT 101 101 Missing (in Ref. 1).  
 SQ SEQUENCE 164 AA; 18696 MW; 9408260374856EE9 CRC64;  
 Query Match 25.3%; Score 618; DB 1; Length 164;  
 Best Local Similarity 84.9%; Pred. No. 11e-37;  
 Matches 118; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

DB 26 LLDPRCLCYLIDGILFIYGVIIITLFLRKFSRSADAPAYOGQOLVNELNLGRREYDV 85  
 QY 383 LKRRADPEWGGQQRNRPOEGVYNALOKDKMAEYSEIGTGERRGGHOSLYGL 442  
 DB 86 LDKRRGRDPEWGGQQRNRPOEGVYNALOKDKMAEYSEIGMGERRGGHOSLYGL 145  
 QY 443 STATRDYDALHMOYLAPR 461  
 DB 146 STATRDYDALHMOYLAPR 164  
 RESULT 12  
 CD3H\_MOUSE STANDARD; PRT; 206 AA.  
 ID CD3H\_MOUSE P20920;  
 AC 01-DEC-1992 (Rel. 24, Created)  
 DT 01-DEC-1992 (Rel. 24, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE T-cell surface glycoprotein CD3 eta chain precursor (T-cell receptor  
 DE T3 eta chain).  
 GN Name=CD3z; Synonyms=CD3h;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Mus.  
 OX NCBI\_Taxid=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A. AND PARTIAL SEQUENCE.  
 RX MEDLINE=90239005; PubMed=2139725;  
 RA Jin Y.J., Clayton L.K., Howard F.D., Koyasu S., Sieh M.,  
 RA Steinblich R., Tarr G.E., Reinherz E.L.;  
 RT "Molecular cloning of the CD3 eta subunit identifies a CD3 zeta-  
 RT related product in thymus-derived cells";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323(1990).  
 RN [2]  
 RP SEQUENCE OF 144-206 FROM N.A.  
 RX MEDLINE=91271358; PubMed=182894;  
 RA Clayton L.K., D'Adamo L., Sieh M., Husey R.E., Koyasu S.,  
 RA Reinherz E.L., Howard F.B.;  
 RT "CD3 eta and CD3 zeta are alternatively spliced products of a common  
 RT genetic locus and are transcriptionally and/or post-transcriptionally  
 RT regulated during T-cell development.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 88:5202-5206(1991).  
 RN [3]  
 RP SEQUENCE OF 144-206 FROM N.A.  
 RX PubMed=2150596;  
 RA Ohno H., Salto T.;  
 RT "CD3 zeta and eta chains are produced by alternative splicing from a  
 RT common gene";  
 RL Int. Immunol. 2:1117-1119(1990).  
 RN [4]  
 RP ERRATUM.  
 RA Ohno H., Salto T.;  
 RL Int. Immunol. 4:1339-1339(1992).  
 CC -1- FUNCTION: Probable role in assembly and expression of the TCR  
 CC complex as well as signal transduction upon antigen triggering.  
 CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a  
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the  
 CC cell surface with the invariant subunits of CD3 labeled gamma,  
 CC delta, epsilon, zeta, and eta. CD3-eta can be complexed in a  
 CC heterodimeric form with CD3-zeta subunit. CD3-eta homodimer has  
 CC not been observed.  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- ALTERNATIVE PRODUCTS:  
 CC Event=Alternative splicing; Named isoforms=2;  
 CC Name=CD-3-eta;  
 CC IsoId=P29020-1; Sequence=Displayed;  
 CC Name=CD-3-zeta;  
 CC IsoId=P24161-1; Sequence=External;  
 CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.  
 CC -1- SIMILARITY: Contains 3 ITAM domains.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration

CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; M3158; AAA37398.1; -  
 DR EMBL; M76711; AAA40403.1; -  
 DR PIR; A35900; A35900.  
 DR MGD; MGI:98334; CD3Z.  
 DR InterPro; IPR003110; ITAM.  
 DR Pfam; PF02189; ITAM; 2.  
 DR SMART; SM00077; ITAM; 2.  
 KW Alternative splicing; Direct protein sequencing; Receptor; Repeat;  
 KW Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 21  
 FT CHAIN 22 206  
 FT DOMAIN 22 30 T-cell surface glycoprotein CD3 eta  
 FT TRANSMEM 31 51 chain.  
 FT DOMAIN 52 206 Extracellular (Potential).  
 FT DOMAIN 69 89 Potential.  
 FT DOMAIN 108 129 Cytoplasmic (Potential).  
 FT DOMAIN 139 159 ITAM 1.  
 FT DISULFID 32 32 ITAM 2.  
 FT DISULFID 32 32 ITAM 3.  
 SQ SEQUENCE 206 AA; 23339 MW; 829256A2CF4E444 CRC64;

Query Match 24.8%; Score 605; DB 1; Length 206;  
 Best Local Similarity 99.2%; Pred. No. 1,4e-36;  
 Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 323 LEDPKCYLLDGLIFLYGVITLALYRAKFSRSASATAANLQDPNOLYNEINIGRREYDV 382  
 DB 26 LIDPKCYLLDGLIFLYGVITLALYRAKFSRSASATAANLQDPNOLYNEINIGRREYDV 85  
 QY 383 LEKRRAPDEMGKQKQRRNPQEGVYNALQDKMAEAYSEIGTGERRRGKHGDLGYQ 440  
 DB 86 LEKRRAPDEMGKQKQRRNPQEGVYNALQDKMAEAYSEIGTGERRRGKHGDLGYQ 143

RESULT 13  
 CD3Z\_PIG STANDARD; PRT; 163 AA.

AC O9XST9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor  
 DE T3 zeta chain).  
 GN Name=CD3Z; (Pig).  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Cetartiodactyla; Suidae; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Minnesota miniature swine;  
 RA Jie H.-B., Yim D., Kim Y.B.;  
 RT "The molecular cloning of porcine CD3 zeta.";  
 RT Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Probable role in assembly and expression of the TCR  
 CC complex as well as signal transduction upon antigen triggering.  
 CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a  
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the  
 CC cell surface with the invariant subunits of CD3 labeled gamma,  
 CC delta, epsilon, zeta, and eta (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor  
 CC triggering (By similarity).  
 CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.  
 CC -1- SIMILARITY: Contains 3 ITAM domains.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

DR EMBL; AF153830; AAD34640.1; -  
 DR InterPro; IPR003110; ITAM.  
 DR Pfam; PF02189; ITAM; 3.  
 DR SMART; SM00077; ITAM; 3.  
 KW Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.  
 FT SIGNAL 1 21  
 FT CHAIN 22 163  
 FT DOMAIN 22 30 T-cell surface glycoprotein CD3 zeta  
 FT TRANSMEM 31 51 chain.  
 FT DOMAIN 52 163 Extracellular (Potential).  
 FT DOMAIN 69 89 Potential.  
 FT DOMAIN 107 128 Cytoplasmic (Potential).  
 FT DOMAIN 138 158 ITAM 1.  
 FT DISULFID 32 32 ITAM 2.  
 FT MOD RES 152 152 ITAM 3.  
 SQ SEQUENCE 163 AA; 18568 MW; 34898620B67167C7 CRC64;

Query Match 24.7%; Score 602.5; DB 1; Length 163;  
 Best Local Similarity 84.2%; Pred. No. 1.6e-36;  
 Matches 117; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY 323 LEDPKCYLLDGLIFLYGVITLALYRAKFSRSASATAANLQDPNOLYNEINIGRREYDV 382  
 DB 26 LIDPKCYLLDGLIFLYGVITLALYRAKFSRSASATAANLQDPNOLYNEINIGRREYDV 85  
 QY 383 LEKRRAPDEMGKQKQRRNPQEGVYNALQDKMAEAYSEIGTGERRRGKHGDLGYQ 442  
 DB 86 LIDRRRGDPEMGK -PRRKQPGGLYNEIQDKMAEAYSEIGKGERRRGKHGDLGYQ 144  
 QY 443 STATKDTYALHMOQLAPR 461  
 DB 145 STATKDTYALHMOQLAPR 163

RESULT 14  
 CD3Z\_RABIT STANDARD; PRT; 165 AA.

AC O9TUF8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor  
 DE T3 zeta chain).  
 GN Name=CD3Z;  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B/J X Chdb:HM;  
 RA Isono T., Nishimura M.;  
 RT "Rabbit CD3 zeta.";  
 RT Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.  
 CC -1- FUNCTION: Probable role in assembly and expression of the TCR  
 CC complex as well as signal transduction upon antigen triggering.  
 CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a  
 CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the  
 CC cell surface with the invariant subunits of CD3 labeled gamma,  
 CC delta, epsilon, zeta, and eta (By similarity).  
 CC -1- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor  
 CC triggering (By similarity).  
 CC -1- SIMILARITY: Belongs to the CD3Z/FCER1G family.



Query Match	24.1%	Score 589.5	DB 1	length 155
Best Local Similarity	81.6%	Pred No. 1.4e-35		
Matches 115	Conservative 9	Mismatches 14	Indels 3	Gaps 2

Qy	323	LEDRKCYLJLGGILFIYVIVITATLRKPKFSRSASFTAAILOPNOLYNEIILNGREEBYV	382
Db	26	LDPRKCYLJLGGILFIYVIVITATLRKPKFSRSASFTAAILOPNOLYNEIILNGREEBYV	85
Qy	383	LEKRRADPEWGGKQORRRNPQEGVYNALQDKMAEAYSSEITKGB--RRRGKGDGLYQ	440
Db	86	LDKRRGDPEWGGK-QRRKNPQEGVYNALQDKMAEAYSSEIMKGNQORRRRGKGDGLYQ	144
Qy	441	GLSTATKDTYDALHMOTLAPR 461	
Db	145	GLSATKDTYDALHMOTLPPR 165	
RESULT 15			
Q99LA6			
ID	Q99LA6	PRELIMINARY;	PRT; 484 AA.
AC	Q99LA6;		
DT	01-JUN-2001 (TREMBlrel. 17, Created)		
DT	01-JUN-2001 (TREMBlrel. 17, Last sequence update)		
DT	01-MAR-2004 (TREMBlrel. 26, Last annotation update)		
DE	Hypothetical protein.		
OS	Mus musculus (Mouse).		
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
ON	NCBI_TaxId=10090;		
RX	[1]		
RP	SEQUENCE FROM N.A.		
RC	STRAIN=mx FVB/N; TISSUE=Mammary tumor;		
RC	MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;		
RA	Strausberg R.J., Feingold E.A., Grouse L.H., Derge J.G.,		
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,		
RA	Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,		
RA	Hopkins R.F., Jordan H., Moore T., Wax S.I., Wang J., Hsieh F.,		
RA	Diatchenko L., Maruina K., Farmer A.A., Rubin G.W., Hong L.,		
RA	Scapleton M., Soares M.B., Bonaldi M.F., Casavant T.L., Scheetz T.E.,		
RA	Brownstein M.J., Uebli T.B., Toshiyuki S., Carninci P., Prange C.,		
RA	Raha S.S., Loguanello N.A., Peters G.J., Abramson R.D., Mullany S.J.,		
RA	Bosak S., McGean P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,		
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,		
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,		
RA	Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,		

RA Whiting M., Madsen A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,  
 RA Krzywinski M.I., Skalska U., Smallus D.B., Schnerch A., Schein J.E.,  
 RA Jones S.J., Mair M.A.?  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 RT and mouse cDNA sequences."?  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
 RP [2].  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=mix FVB/N; TISSUE=Mammary tumor;  
 RA Strausberg R.;  
 RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.  
 DR EMBL: BC003495; AA0103495.1; -.  
 DR PIR: F33932; F33932.  
 DR PIR: S19963; S19963.  
 DR HSPB; P01810; 2FBJ.  
 DR InterPro: IPR007110; IG-like.  
 DR InterPro: IPR003597; IG\_c1.  
 DR InterPro: IPR003006; IG\_MHC.  
 DR InterPro: IPR003596; IG\_v.  
 DR Pfam: PF07654; C1-sect; 2.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PSS0835; IG LIKE; 4.  
 DR PROSITE; PS00290; IG\_MHC; UNKNOWN\_2.  
 DR Hypothetical protein.  
 QO SEQUENCE 484 AA; 52567 MW; 8EA4E4FBCF582FA CRC64;

Query Match Similarity	23.3%	Score	569	DB 2	Length	484
Best Local Similarity	40.9%	Pred. No.	1.8e-33			
Matches 157	Conservative 36	Mismatches 79	Indels 112	Gaps 19		

[illegible]

Search completed: April 20, 2005, 07:07:08  
Job time : 68.7843 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:48:35 ; Search time 65.1759 Seconds  
(without alignments)  
2628.607 Million cell updates/sec

Title: US-09-596-774-7

Perfect score: 2345  
Sequence: 1 QIQVQSGPELKKPGETVKI.....LSTATKDYDALHMQTLAPR 443

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_16Dec04:\*

- 1: geneseqp1980s:\*
- 2: geneseqp1990s:\*
- 3: geneseqp2000s:\*
- 4: geneseqp2001s:\*
- 5: geneseqp2002s:\*
- 6: geneseqp2003as:\*
- 7: geneseqp2003bs:\*
- 8: geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2345	100.0	461	2	AAW85508
2	1636.5	69.8	483	2	AAW82315
3	1501.5	64.0	428	2	AAW24027
4	1501.5	64.0	443	2	AAW24025
5	1308	55.8	615	2	AAW95056
6	1308	55.8	617	2	AAW95057
7	1303	55.6	711	2	AAW26980
8	1302	55.5	530	2	AAW95053
9	1302	55.5	651	2	AAW05136
10	1302	55.5	699	2	AAW05138
11	1302	55.5	892	2	AAW05140
12	1302	55.5	892	2	AAW05143
13	1302	55.5	892	2	AAW05139
14	1302	55.5	895	2	AAW05142
15	1302	55.5	899	2	AAW05144
16	1302	55.5	1020	2	AAW05141
17	1299	55.4	637	2	AAW26982
18	1284	54.8	240	2	AAW85494
19	1284	54.8	240	2	AAW05134
20	1275	54.4	241	2	AAW26979
21	1249	53.3	631	7	ADH34596
22	1015.5	43.3	443	6	ABG76488
23	1015.5	43.3	443	6	ABG74240
24	1013	43.2	242	4	AAW20433
25	1011.5	43.1	473	2	AAW26646

26	1011.5	43.1	514	2	AAW26647	AAW26647	Chimeric
27	981.5	41.9	250	2	AAW02278	AAW02278	741F8 ant
28	981.5	41.9	250	2	AAW29261	AAW29261	Anti-C-er
29	981.5	41.9	250	2	AAW22400	AAW22400	Single ch
30	981.5	41.9	250	2	AAW47012	AAW47012	Single ch
31	981.5	41.9	250	2	AAW53168	AAW53168	741F8 ant
32	981.5	41.9	250	2	AAW80422	AAW80422	Anti-C-er
33	981.5	41.9	250	7	ABW00714	ABW00714	8FV
34	975.5	41.6	259	7	ADG32322	ADG32322	Mouse scf
35	975.5	41.6	329	7	ADG32359	ADG32359	Mouse scf
36	963.5	41.1	667	2	AAW39573	AAW39573	Sequence
37	962.5	41.0	651	2	AAW26649	AAW26649	Chimeric
38	962.5	41.0	692	2	AAW26650	AAW26650	Chimeric
39	960.5	41.0	259	7	ADG32324	ADG32324	Mouse scf
40	960.5	41.0	329	7	ADG32361	ADG32361	Precursor
41	930	39.7	444	2	AAW36845	AAW36845	Single ch
42	921	39.3	240	8	ADN16154	ADN16154	Mouse C7
43	917.5	39.1	505	5	AAU72875	AAU72875	Human NKG
44	917	39.1	251	7	ADG32321	ADG32321	Mouse scf
45	917	39.1	320	7	ADG32358	ADG32358	Precursor

## ALIGNMENTS

RESULT 1									
AAW85508	ID	AAW85508	standard; protein; 461 AA.						
XX	XX	AAW85508;							
XX	XX	16-MAR-1996	(first entry)						
DT	DT								
XX	XX								
DE	DE	Leader-scFv(FRP5):lyt-2 hinge;zeta.							
XX	XX								
KW	KW	Single chain antibody; scFv; antibody engineering; antitumour;							
KW	KW	tumour antigen binding; T-cell receptor; cytotoxic T-lymphocyte;							
KW	KW	monoclonal antibody; erbB-2; cancer; cell targeting;							
KW	KW	adoptive immunotherapy.							
OS	OS	Synthetic.							
XX	XX								
XX	XX	Key	Location/Qualifiers						
FT	FT	Peptide	40..93						
FT	FT	Domain	/label= IGH_chain_leader						
FT	FT	Peptide	94..819						
FT	FT	Peptide	/label= scFv(FRP)5						
FT	FT	Domain	261..322						
FT	FT	Domain	/label= lyt-2_hinge						
FT	FT	Domain	323..461						
FT	FT	Domain	/label= zeta_chain						
XX	XX								
XX	XX	WO9530014-A1.							
XX	XX	09-NOV-1995.							
XX	XX	20-APR-1995;	95WO-EP001494.						
XX	XX	02-MAY-1994;	94EP-00810244.						
XX	XX	(CIBA ) CIBA GEIGY AG.							
XX	XX	Groner B, Moritz D;							
XX	XX	WPI; 1995-393085/50.							
XX	XX	N-PSDB; AAT05783.							
XX	XX								
XX	XX	New bifunctional proteins for use in killing tumour cells - contg. a							
XX	XX	tumour antigen binding domain, a hinge region and a zeta chain derived							
XX	XX	from a T-cell antigen receptor.							
XX	XX	Example 1; Page 29-31; 46pp; English.							

CC A bifunctional protein (AAR85505) consists of a single chain antibody, scFv(RRP)5, directed against the tumour erbB-2 antigen, a hinge region, CC and a functional zeta chain obtd. from a T-cell receptor. The protein is CC expressed in host cells, esp. cytotoxic T-lymphocytes, providing them CC with a defined tumour cell specificity enabling targeting to defined CC tumour cells and MHC-unrestricted and MHC-independent tumour destruction CC in vitro or in vivo

XX Sequence 461 AA;

Query Match 100.0%; Score 2345; DB 2; Length 461;

Best Local Similarity 100.0%; Pred. No. 1,2e-155; Mismatches 0; Indels 0; Gaps 0;

Matches 443; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 QIOLVSGPELKKRGETVTKISCKASGYPTNYGMWVQAQGGKMKMGWINTSTGESETF 60  
 DB 19 QIOLVSGPELKKRGETVTKISCKASGYPTNYGMWVQAQGGKMKMGWINTSTGESETF 78  
 QY 61 ADDPKGRFDSLETSANTAYLQINNLSKEDMATYFCARMEYVHGVPYWGQTTVTYSSG 120  
 DB 79 ADDPKGRFDSLETSANTAYLQINNLSKEDMATYFCARMEYVHGVPYWGQTTVTYSSG 138  
 QY 121 GGGSGGGGGGGSDIQLTQSHKFLSTSVGRVSITCKASQDVYNAVAMYOQKQSPKL 180  
 DB 139 GGGSGGGGGGGSDIQLTQSHKFLSTSVGRVSITCKASQDVYNAVAMYOQKQSPKL 198  
 QY 181 LIYASASSRYTVPSRFTSGSGSPDFTTSSVOAEDLAVYCCQHFRTPTFGSGTKLEI 240  
 DB 199 LIYASASSRYTVPSRFTSGSGSPDFTTSSVOAEDLAVYCCQHFRTPTFGSGTKLEI 258  
 QY 241 KALEISNSVMYFSSVVPVLOKNSVTTTPVLRTPSPVHPPTGSOPOREDCRPGSVKGT 300  
 DB 259 KALEISNSVMYFSSVVPVLOKNSVTTTPVLRTPSPVHPPTGSOPOREDCRPGSVKGT 318  
 QY 301 GLDFLEDPKLCYLLDGLIFIVGVIIITALLYLRAKFSRSAETAANLQDPNQLYNEINLGRRE 360  
 DB 319 GLDFLEDPKLCYLLDGLIFIVGVIIITALLYLRAKFSRSAETAANLQDPNQLYNEINLGRRE 378  
 QY 361 EYDVLKRRADPEWGGKQRRRNPOEGVYNALQDKMAEAYSEIGTGERRRKGHDGL 420  
 DB 379 EYDVLKRRADPEWGGKQRRRNPOEGVYNALQDKMAEAYSEIGTGERRRKGHDGL 438  
 QY 421 YQGLSTATKDTYDALHMOTLAPR 443  
 DB 439 YQGLSTATKDTYDALHMOTLAPR 461

RESULT 2

AAM82315 ID AAM82315 standard; protein; 483 AA.

XX AAM82315;

XX 26-FEB-1999 (first entry)

XX Chimeric CD19/CTCR protein.

XX Single chain; antibody; B-cell marker; CD19; ScFvCD19; gene therapy;  
 XX retroviral vector; tumour-associated antigen; cancer; immunoglobulin;  
 XX CD8 alpha chain; hinge region; beta chain; transmembrane domain;  
 XX intracellular domain; T-cell receptor-CD3 complex.

OS Homo sapiens.

OS Synthetic.

PN DE19720152-A1.

PD 05-NOV-1998.

PF 02-MAY-1997; 97DE-01020152.

PR 02-MAY-1997; 97DE-01020152.

PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.

XX Schroeder A, Schwenkenbecher J, Bargou R;

XX WPI; 1998-584686/50.

DR N-PSDB; AAV73324.

PT Retroviral vector system for gene therapy - by expression of therapeutic gene in T cells.

PS Claim 9; Page 6; 10pp; German.

CC This sequence represents a novel chimeric CD19/CTCR protein which is  
 CC composed of an immunoglobulin leader sequence, an anti-CD19 single chain  
 CC antibody fragment (ScFvCD19), a hinge region from the CD8 alpha chain and  
 CC the transmembrane and intracellular domain of the T-cell receptor-CD3  
 CC complex zeta chain. This sequence is used in a retroviral vector system  
 CC for gene transfer and expression in primary human T cells which involves  
 CC a retroviral vector containing CDNA corresponding to a therapeutic gene,  
 CC a promoter selected from the SV40 immediate early promoter or the human  
 CC II-2 or MHC I promoter and a selectable marker, and a packaging cell line  
 CC (i.e. a modified retroviral packaging line for enhanced transfer into  
 CC primary human T lymphocytes). This vector system in which the CDNA  
 CC encodes a single-chain antibody (scFv) specific for a tumour-associated  
 CC antigen can be used for gene therapy of cancer, especially by isolating T  
 CC cells from the patient's blood, stimulating the T cells in vitro,  
 CC introducing the retroviral vector into the T cells and returning the T  
 CC cells to the patient, preferably by intravenous injection

XX Sequence 483 AA;

Query Match 69.8%; Score 1636.5; DB 2; Length 483;

Best Local Similarity 70.3%; Pred. No. 4,2e-106; Mismatches 326; Conservative 36; Mismatches 79; Indels 21; Gaps 5;

QY 1 QIOLVSGPELKKRGETVTKISCKASGYPTNYGMWVQAQGGKMKMGWINTSTGESETF 60  
 DB 20 EVQLQESGAEIVRSGSVKISCKASGYAFSSWMWVKRPGQGLIEWIGIWDGDTNY 79  
 QY 61 ADDPKGRFDSLETSANTAYLQINNLSKEDMATYFCARMEV-----HGVVVPYWGQTTV 115  
 DB 80 NGKPKGRATYLADESSSTAAMWQLSLASEDSAVYFCARRRTTVGVPPYVAMDWQGSTSC 139  
 QY 116 TVSSGGGGS--GGGSGGGSDIQLTQSHKFLSTSVGRVSITCKASQDY--VAV 169  
 DB 140 TASAKTTPKLBEBEPEARVDILITQSPASLAVALQGRATISCKASQSDYDSDSLNW 199  
 QY 170 YQKPGSPKLLIYASASSRYTVPSRFTSGSGSPDFTTSSVOAEDLAVYCCQHFRTPT 229  
 DB 200 YQIIPGQPKLLIYDASNLVSGIIPRPSGSGSGDTFLNIHPEKVDAAATYHCQOSTEDP 259  
 QY 230 FTSSGSKLEIKA-----LEISNSVMYFSSVVPVLOKNSVTTTPVLRTPSPVHP 279  
 DB 260 WTFGGGTRKLEILEQKLISEDLNLEISNSVMYFSSVVPVLOKNSVTTTPVLRTPSPVHP 319  
 QY 280 TGTSPQRPEDCRPRGSGVKGTLDFLEDPKLCYLLDGLIFIVGVIIITALLYLRAKFSRSAE 339  
 DB 320 TGTSPQRPEDCRPRGSGVKGTLDFLEDPKLCYLLDGLIFIVGVIIITALLYLRAKFSRSAE 379  
 QY 340 TAANLQDPNQLYNEINLGRREYDVLKRRADPEWGGKQRRRNPOEGVYNALQDKMA 399  
 DB 380 TAANLQDPNQLYNEINLGRREYDVLKRRADPEWGGKQRRRNPOEGVYNALQDKMA 439  
 QY 400 EAYSEIGTGERRRKGHDGLYQGLSTATKDTYDALHMOTLAPR 443  
 DB 440 EAYSEIGTGERRRKGHDGLYQGLSTATKDTYDALHMOTLAPR 483

RESULT 3

AAM24027 ID AAM24027 standard; protein; 428 AA.

XX AAM24027;

XX 25-MAR-2003 (revised)  
 DT 04-MAR-1998 (first entry)  
 XX Single chain antigen hybrid receptor partial sequence.  
 DE Hybrid receptor; single chain antigen; gene therapy; diagnosis;  
 KW signal conduction; receptor; control region.  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FT Peptide 1..4  
 FT /note= "partial leader sequence"  
 FT Protein 5..428  
 FT /note= "single chain antigen hybrid receptor"  
 FT Region 5..124  
 FT /label= VH\_B\_1-8  
 FT /note= "variable heavy chain region of B1-8 antibody"  
 FT Region 125..139  
 FT /label= (GLY\_Ser)3  
 FT /note= "linker region"  
 FT Region 140..249  
 FT /label= VL\_B\_1-8  
 FT /note= "variable light chain region of B1-8 antibody"  
 FT Region 250..294  
 FT /label= CD\_8alpha\_hinge  
 FT Domain 295..315  
 FT /label= transmembrane\_domain  
 FT Domain 331..428  
 FT /label= cytoplasmic\_domain  
 XX  
 PN W09720938-A2.  
 XX  
 XX 12-JUN-1997.  
 PD  
 XX 03-DEC-1996; 96WO-DE002334.  
 PF  
 XX 05-DEC-1995; 95DE-01045351.  
 PR  
 XX (UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.  
 PA  
 XX Mettelmann R, Kulmburg P, Rosenthal F;  
 PI  
 XX WPI; 1997-319784/29.  
 DR N-PSDB; AAT77139.  
 DR  
 XX  
 XX Cells with hybrid receptor having extracellular and intracellular regions  
 PT of different origins - useful in gene therapy and diagnosis of tumours.  
 PT  
 XX Example 3; Fig 4; 46pp; German.  
 PS  
 XX  
 CC This partial sequence encodes a novel single chain antigen hybrid  
 CC receptor (HR). The full length receptor sequence can be found in AAM24025  
 CC and contains an extracellular domain specific for the hapten 4-hydroxy-5-  
 CC 1000-3-nitrophenyl acetate (NIP) coupled to the CD8-alpha region and the  
 CC transmembrane and signal-conducting intracellular parts of the CD3-zeta  
 CC molecule. Such hybrid receptors comprise a receptor part localised on the  
 CC outside of the cell and specific to a particular signal molecule and a  
 CC receptor part originating from another receptor, localised on the inside  
 CC of the cell and capable of setting off a signal inside the cell. The cell  
 CC should also contain at least one other gene construct with a control  
 CC region which can interact with the signal sent out by the hybrid receptor  
 CC and thereby control expression of a transgene bound to this control  
 CC region. Such cells are useful in gene therapy or for diagnostic purposes.  
 CC (Updated on 25-MAR-2003 to correct PR field.)  
 CC  
 XX  
 SQ Sequence 428 AA;

QY 1 QIOLVOSGPELKKPEETVYKISCKASGYPTNYGMNVKAPGQGLKMMGWINTSTGESTP 60  
 DB 5 QVQLQSSAEELVYKCAVSKLCKASGTFITVMMHWKQGRGLEGWIGRIDPNSGRKY 64  
 QY 61 ADFKGRPFDSLETSANTAYLIQINNLSKEDMATYPCAREVY-HQYVYMGQGTTVVSS 119  
 DB 65 NEKFSKATLTVDRPSSSTAYVQLSLTSEDAVYVCARYDYSSYFDYMGQGTTVVSS 124  
 QY 120 GGGSGGGSGGGSGGSDIQLTSHKFLSTSVGDRYSITCKAGDYY--NAAVYQKRGQ 176  
 DB 125 GGGSGGGSGGGSGGSAVGTQ-ESALTTSPGETVTLTGRSSGAVTTSNVANWQEKPDH 183  
 QY 177 SPKLTIYASRRYGVPERFTSGSGSPDFTTISVQAEADAVYFCQOHFPTPPFGSGT 236  
 DB 184 LFTGLIGTNNRAAGVPRFSGSLIGDKAALTITGACQTEDAIYFCALMTYNNHWFGGCT 243  
 QY 237 KLEIKALEISNSVMYFSSVPELQKVNSTTKPVLRTPSPVHPGTSGPQRPEDCRPRGS 296  
 DB 244 KLTV--LE-----EFTTKPVLRTPSPVHPGTSGPQRPEDCRPRGS 283  
 QY 297 VKGTGLDFLEDPKLCYLLDGLFTYGVITTLTYLPAKRSRAETNANLQDPNLYENLT 356  
 DB 284 VKGTGLDF--DPKLCYLLDGLFTYGVITTLTYLRRKFSRAETNANLQDPNLYENLT 341  
 QY 357 GRREYDYLEKKRRARDPEMGKQQRNRRPOEGVYNALQKDAEAYSEITGKERRRGK 416  
 DB 342 GRREYDYLEKKRRARDPEMGKQQRNRRPOEGVYNALQKDAEAYSEITGKERRRGK 401  
 QY 417 HDGLYQGLSTATKDTYDALHMOTLAPR 443  
 DB 402 HDGLYQGLSTATKDTYDALHMOTLAPR 428

RESULT 4  
 AAM24025  
 ID AAM24025 standard; protein; 443 AA.  
 XX  
 XX AAM24025;  
 AC  
 XX 25-MAR-2003 (revised)  
 DT 04-MAR-1998 (first entry)  
 DT  
 XX Single chain antigen hybrid receptor.  
 DE  
 XX Hybrid receptor; single chain antigen; gene therapy; diagnosis;  
 KW signal conduction; receptor; control region.  
 KW  
 XX  
 OS Synthetic.  
 OS Homo sapiens.  
 XX  
 XX Key Location/Qualifiers  
 FH Peptide 1..19  
 FT /label= leader\_peptide  
 FT Protein 20..444  
 FT /note= "single chain antigen hybrid receptor"  
 FT Region 20..139  
 FT /label= VH\_B\_1-8  
 FT /note= "variable heavy chain region of B1-8 antibody"  
 FT Region 140..154  
 FT /label= (GLY\_Ser)3  
 FT /note= "linker region"  
 FT Region 155..264  
 FT /label= VL\_B\_1-8  
 FT /note= "variable light chain region of B1-8 antibody"  
 FT Region 265..309  
 FT /label= CD\_8-alpha\_hinge  
 FT Domain 310..330  
 FT /label= transmembrane\_domain  
 FT Domain 331..444  
 FT /label= cytoplasmic\_domain  
 FT  
 XX  
 PN W09720938-A2.



Key	Location/Qualifiers
FT Peptide	1. .21
FT	/label= ompa_signal_peptide
FT Domain	23. .141

FT	Peptide	/label= FRP5_heavy_chain_variable_domain
FT		142..156
FT	Domain	/label= Linker
FT		157..264
FT	Protein	/label= FRP5_light_chain_variable_domain
FT		265..711
XX		/label= phoa
XX	EP502812-A1.	
XX		
PD	09-SEP-1992.	
XX		
PF	27-JAN-1992;	92EP-00810056.
XX		
PR	05-FEB-1991;	91EP-00810079.
XX		
PA	(CIBA ) CIBA GEIGY AG.	
XX		
PI	Wels WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M,	
DR	WPI; 1992-302096/37.	
DR	N-PSDB; AAQ28257.	
PT	Recombinant antibodies directed to growth factor receptor C-erbB-2 - for	
PT	diagnosing and treating tumours expressing C-erbB-2 e.g. breast or	
PT	ovarian tumours.	
XX		
PS	Disclosure; Page 34-40; 67pp; English.	
XX		
CC	The sequence given is the single chain recombinant antibody desiganted	
CC	Fv(FRP5)-phoa. The alkaline phosphatase gene (phoa) was used as a marker	
CC	gene so that E. coli transformed with the fusion gene could be	
CC	identified. The fusion gene was expressed in E. coli and the antibody was	
CC	extracted. This recombinant antibody can be used for the qualitative and	
CC	quantitative determination of c-erbB-2. This can be used for monitoring	
CC	or in-vivo localisation of tumours overexpressing c-erbB-2. (Updated on	
XX	25-MAR-2003 to correct PN field.)	
XX		
XX	Sequence 711 AA;	
XX		

Query Match	55.6%	Score 1303	DB 2	Length 711
Best Local Similarity	94.2%	Pred. No. 1.3e-82		
Matches 245	Conservative 2	Mismatches 1	Indels 12	Gaps 1
QY	1	QIQIVQSGBELKKPGETVAKISCKASGYPTNYGNMNVKQAPQGLKMGMINSTGSEYF	60	
Db	23	QVQLQQSGBELKKPKPEITVAKISCKASGYPTNYGNMNVKQAPQGLKMGMINSTGSEYF	82	
QY	61	ADDFKGRFDFSLSETANTAYILQINNLKSEDMATYCARREYVHGVPYVGQSTTYTVSSG	120	
Db	83	ADDFKGRDFSLSETANTAYILQINNLKSEDMATYCARREYVHGVPYVGQSTTYTVSSG	142	
QY	121	GGGSGGGSGGGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVNVAVMYQOKPQGSPL	180	
Db	143	GGGSGGGSGGGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVNVAVMYQOKPQGSPL	202	
QY	181	LIYSASSRYTVGPSRFETGSGSPDPTFTISSVQADLAVYRCQHFRTPTFGSGTKLEI	240	
Db	203	LIYSASSRYTVGPSRFETGSGSPDPTFTISSVQADLAVYRCQHFRTPTFGSGTKLEI	262	
QY	241	KALEISNSVMYFSSVVPVLQ 260		
Db	263	KALE-----PVLE 270		
RESULT 8				
AAAR95053				
ID	AAAR95053	standard; protein; 530 AA.		
XX	AAAR95053;			
XX				
DT	16-OCT-2003	(revised)		
DT	18-AUG-1996	(first entry)		

XX		scFv(FRP5) - DELTA-DGAL4 multidomain protein.
DE		
KW	Nucleic acid transfer system; gene transfer; gene therapy;	
KX	cell targeting; multidomain protein; vector; cancer; exotoxin A; DELTA,	
XX	single chain antibody; scFv; Gal4.	
OS	Mus; sp.	
OS	Pseudomonas; aeruginosa.	
OS	Saccharomyces cerevisiae.	
OS	Chimeric.	
OS	Synthetic.	
PH		
FH	Key	Location/Qualifiers
FT	Peptide	1..8
FT		/label= FLAG_epitope
FT	Peptide	9..17
FT		/label= Spacer
FT	Domain	18..257
FT		/label= ScFv(FRP5)
FT	Peptide	258..260
FT		/label= Spacer
FT	Domain	261..375
FT		/label= ERA
FT	Peptide	/note= "amino acids 252-366 of exotoxin-A"
FT		376
FT		/label= Spacer
FT	Domain	377..522
FT		/label= GAL4
FT	Peptide	/note= "amino acids 2-147 of yeast GAL4"
FT		523..530
FT		/label= Spacer
FT		/note= "endoplasmic reticulum retention peptide"
PN		
XN	W09613599-A1.	
PD		
XX	09-MAY-1996.	
XX		
PF	31-OCT-1995; 95WO-EP004270.	
XX		
PR	01-NOV-1994; 94EP-00810627.	
PA	(WELS/) WELS W.	
XX		
PI	Weis W, Fominaya J;	
XX		
DR	WPI, 1996-239505/24.	
DR	N-PSDB; AAT29409.	
XX		
PT	Nucleic acid transfer system for gene therapy, e.g. against cancer -	
PT	includes toxin translocation domain to target nucleic acid to specific	
XX	cell.	
BS	Claim 7; Page 59-61; 106pp; English.	
CC	A multidomain protein (AAR35053) has a FLAG epitope, a single chain	
CC	antibody, scFv, of monoclonal antibody FRP5 (raised against human tumour	
CC	cell HER2 antigen) that acts as a ligand domain, a non-cytotoxic portico	
CC	of Pseudomonas aeruginosa exotoxin A acting as a translocation domain a	
CC	the DNA binding domain of yeast GAL4. It is the product of a fusion gen	
CC	(AAT29409) and can be expressed in E. coli (resulting in removal of omp	
CC	signal peptide). It is used with an effector nucleic acid that comprise	
CC	e.g. a gene to be delivered to a cell and a cognate structure for the	
CC	GAL4 DNA binding domain. This provides a novel means of nucleic acid	
CC	transfer, suitable for gene therapy. (Updated on 16-Oct-2003 to	
XX	standardise OS field)	
XX		
Sequence 530 AA;		
50		

```

Query Match      55.5%;   Score 1302;   DB 2;   Length 530;
Best Local Similarity 99.2%;   Pred. No. 1.1e-82;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```



```

QY 1 QIOLVQSGPELKKPGETYKISCKASGYPTNYGMNWKQAPQGGLKMMGMINSTGSESTF 60
DB 18 QVQLQSGPELKKPGETYKISCKASGYPTNYGMNWKQAPQGGLKMMGMINSTGSESTF 77
QY 61 ADDPKGRDFEISLETSANTAYLIQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVSSG 120
DB 78 ADDPKGRDFEISLETSANTAYLIQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVSSG 137
QY 121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQOKPGQSPKL 180
DB 138 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQOKPGQSPKL 197
QY 181 LIYSASRYTGVPSRFITGSGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 240
DB 198 LIYSASRYTGVPSRFITGSGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 257
QY 241 KALE 244
DB 258 KALE 261

```

## RESULT 9

AAW05136  
ID AAW05136 standard; protein; 651 AA.

AC AAW05136;

DT 29-JAN-1997 (first entry)

DE scFv(FRP5)-ETA fusion protein.

Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;

epidermal growth factor; receptor; plasmid pSW202-5; cancer; therapy;

antitumor; exotoxin A; ETA.

Mus; BP.

Pseudomonas; aeruginosa.

OS Synthetic.

OS Chimeric.

Key Location/Qualifiers

FT 1..21

FT /label= Sig peptide

FT /note= "ompA signal peptide"

FT 22..38

FT /label= Spacer

FT 39..278

FT /label= scFv(FRP5)

FT 279..289

FT /label= Spacer

FT 290..651

FT /label= ETA

FT /note= "exotoxin A amino acids 252-613"

EP739984-A1.

30-OCT-1996.

26-APR-1995; 95EP-00106275.

26-APR-1995; 95EP-00106275.

(SANT-) SAN TUMORFORSCHUNGS GMBH.

Weis W, Schmidt M, Groner B;

WPI; 1996-478748/48.

N-ESDB; AAT42036.

Bivalent fusion proteins that bind epidermal growth factor receptor or

analogues - and comprise at least two different cell surface binding

domain(s), useful for tumour therapy.

PS Example 7; Page 21-22; 52pp; English.

scFv(FRP5)-ETA (AAW05136) comprises the single-chain binding region (see

also AAW05134) of murine monoclonal antibody FRP5, which is specific for

human epidermal growth factor receptor erbB-2, joined to exotoxin A

(ETA). It is encoded by plasmid pSW202-5 (see also AAT42036) obt'd. by

ligating an scFv(FRP5) gene (AAT42034) into plasmid pSW200 contg. the

Pseudomonas aeruginosa PAK ETA gene. The construct can be used to produce

novel bivalent fusion proteins (see also AAW05135-44) in bacterial host

cells, for use as antitumor agents

Sequence 651 AA;

Query Match 55.5%; Score 1302; DB 2; Length 651;

Best Local Similarity 99.2%; Pred. No. 1,4e-82;

Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

1 QIOLVQSGPELKKPGETYKISCKASGYPTNYGMNWKQAPQGGLKMMGMINSTGSESTF 60

39 QVQLQSGPELKKPGETYKISCKASGYPTNYGMNWKQAPQGGLKMMGMINSTGSESTF 98

61 ADDPKGRDFEISLETSANTAYLIQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVSSG 120

99 ADDPKGRDFEISLETSANTAYLIQINNLSKEDMATYFCARMEVYHGVPYWGQGTITVSSG 158

121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQOKPGQSPKL 180

159 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAMVYQOKPGQSPKL 218

181 LIYSASRYTGVPSRFITGSGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 240

219 LIYSASRYTGVPSRFITGSGSPDFTFTISSVQAEIDLAVYFCQGHFRPTFTFGSGTKLEI 278

241 KALE 244

279 KALE 282

RESULT 10

AAW05138

ID AAW05138 standard; protein; 699 AA.

AC AAW05138;

DT 29-JAN-1997 (first entry)

DE scFv(FRP5)/TGF alpha-ETA.

Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;

epidermal growth factor; receptor; plasmid pMS238-5-TGF; cancer;

exotoxin A; ETA; transforming growth factor alpha; TGF; antitumor.

Homo; sapiens.

Mus; BP.

Pseudomonas; aeruginosa.

OS Synthetic.

OS Chimeric.

Key Location/Qualifiers

FT 1..21

FT /label= Sig peptide

FT /note= "ompA signal peptide"

FT 22..38

FT /label= Spacer

FT 39..278

FT /label= scFv(FRP5)

FT 279..289

FT /label= Spacer

FT 290..404

FT /label= ETA

FT /note= "exotoxin A amino acids 252-366"

FT 405..410

FT /label= Spacer

```

FT Region 411..460
FT /label= TGF-alpha
FT Peptide 461..465
FT /label= Spacer
FT Region 466..699
FT /label= ETA
FT /note= "endotoxin-A amino acids 380-613"
EP739984-A1.
PD 30-OCT-1996.
XX 26-APR-1995; 95EP-00106275.
XX 26-APR-1995; 95EP-00106275.
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX Wels W, Schmidt M, Groner B;
XX MPI, 1996-478748/48.
XX N-PSDB; AAT42038.
DR Bivalent fusion proteins that bind epidermal growth factor receptor or
XX analogues - and comprise at least two different cell surface binding
XX domain(s), useful for tumour therapy.
XX
XX Example 10; Page 25-27; 52pp; English.
XX
XX scFv(FRP5)/TGF alpha-ETA (AAW05138) comprises the single-chain binding
XX region of murine monoclonal antibody FRP5 (specific for human epidermal
XX growth factor receptor erbB-2, see also AAW05134) joined to portions of
XX exotoxin A from Pseudomonas aeruginosa and to human transforming growth
XX factor (TGF) alpha (see also AAW05137). It is encoded by plasmid pMS238-5
XX -TGF (AAT42038). This plasmid can be utilised in the prodn. of the
XX bi-specific fusion protein in bacterial (esp. E. coli) host cells. Such
XX fusion proteins (see also AAW05139-44) are useful as antitumour agents
XX
XX Sequence 699 AA:
SQ
Query Match 55.5%; Score 1302; DB 2; Length 699;
Best Local Similarity 99.2%; Pred. No. 1.5e-82;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 QIQLVSGPELKKKPGETVKISCKASGYPTNYGMNWKQAPQGLKMGWINTSGESTF 60
DB 39 QVQLQSGPELKKKPGETVKISCKASGYPTNYGMNWKQAPQGLKMGWINTSGESTF 98
QY 61 ADDPKGRFDPSLETSANTAYLQINNLKSEDMATYFCARWEVYHGYVPWGGTTVTVSSG 120
DB 99 ADDPKGRFDPSLETSANTAYLQINNLKSEDMATYFCARWEVYHGYVPWGGTTVTVSSG 158
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVNNAVAWYQQRKQSGPKL 180
DB 159 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVNNAVAWYQQRKQSGPKL 218
QY 181 LIYSASSRRTYGVPRFTGSGSGPDFTTISVQAEDLAVVFCQCHFTPTFFSGSTLEI 240
DB 219 LIYSASSRRTYGVPRFTGSGSGPDFTTISVQAEDLAVVFCQCHFTPTFFSGSTLEI 278
QY 241 KALE 244
DB 279 KALE 282
XX
RESULT 11
AAW05140
ID AAW05140 standard; protein; 892 AA.
XX
XX AAW05140;
XX
XX 29-JAN-1997 (first entry)
XX

```

```

DE scFv2(225/FRP5)-ETA.
XX
XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
XX epidermal growth factor; receptor; plasmid pMS238-225-5; cancer;
XX exotoxin A; ETA; antitumour.
XX
XX Mus; gp.
XX Pseudomonas; aeruginosa.
XX Synthetic.
XX Chimeric.
XX
XX Key Location/Qualifiers
XX FH 1..21
XX FH /label= sig_peptide
XX FT /note= "ompA signal peptide"
XX FT 22..38
XX FT /label= Spacer
XX FT 39..278
XX FT /label= scFv(225)
XX FT 279..289
XX FT /label= Spacer
XX FT 290..404
XX FT /label= ETA
XX FT /note= "exotoxin A amino acids 252-366"
XX FT 405..407
XX FT /label= Spacer
XX FT 408..647
XX FT /label= scFv(FRP5)
XX FT 648..658
XX FT /label= Spacer
XX FT 659..892
XX FT /label= ETA
XX /note= "endotoxin-A amino acids 380-613"
XX
XX EP739984-A1.
XX
XX 30-OCT-1996.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX MPI, 1996-478748/48.
XX N-PSDB; AAT42040.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX analogues - and comprise at least two different cell surface binding
XX domain(s), useful for tumour therapy.
XX
XX Example 11; Page 31-33; 52pp; English.
XX
XX scFv2(225/FRP5)-ETA (AAW05140) comprises the single-chain binding region
XX of murine monoclonal antibody 225 (specific for human epidermal growth
XX factor receptor, see also AAW05133) joined to portions of exotoxin A from
XX Pseudomonas aeruginosa and to the single-chain binding region of murine
XX monoclonal antibody FRP5 (specific for human epidermal growth factor
XX receptor erbB-2, see also AAW05134). It is encoded by plasmid pMS238-225-
XX 5 (AAT42040). This plasmid can be utilised in the prodn. of the bivalent
XX fusion protein in bacterial (esp. E. coli) host cells. Such fusion
XX proteins (see also AAW05138-44) are useful as antitumour agents
XX
XX Sequence 892 AA:
SQ
Query Match 55.5%; Score 1302; DB 2; Length 892;
Best Local Similarity 99.2%; Pred. No. 2e-82;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 QIQLVSGPELKKKPGETVKISCKASGYPTNYGMNWKQAPQGLKMGWINTSGESTF 60
DB 39 QVQLQSGPELKKKPGETVKISCKASGYPTNYGMNWKQAPQGLKMGWINTSGESTF 98
QY 61 ADDPKGRFDPSLETSANTAYLQINNLKSEDMATYFCARWEVYHGYVPWGGTTVTVSSG 120
DB 99 ADDPKGRFDPSLETSANTAYLQINNLKSEDMATYFCARWEVYHGYVPWGGTTVTVSSG 158
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVNNAVAWYQQRKQSGPKL 180
DB 159 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVNNAVAWYQQRKQSGPKL 218
QY 181 LIYSASSRRTYGVPRFTGSGSGPDFTTISVQAEDLAVVFCQCHFTPTFFSGSTLEI 240
DB 219 LIYSASSRRTYGVPRFTGSGSGPDFTTISVQAEDLAVVFCQCHFTPTFFSGSTLEI 278
QY 241 KALE 244
DB 279 KALE 282
XX

```



```

FT FT /label= Spacer
FT FT 290..404
FT FT /label= ETA
FT FT /note= "exotoxin A amino acids 252-366"
FT FT Peptide
FT FT 405..407
FT FT /label= Spacer
FT FT Region
FT FT 408..647
FT FT /label= scFv(225)
FT FT Peptide
FT FT 648..658
FT FT /label= Spacer
FT FT Region
FT FT 659..892
FT FT /label= ETA
FT FT /note= "endotoxin-A amino acids 380-613"

XX EP739984-A1.
XX
XX 30-OCT-1996.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX
XX WPI; 1996-478748/48.
XX
XX N-PSDB; AAT42039.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX domain(s) - and comprise at least two different cell surface binding
XX domain(s), useful for tumour therapy.
XX
XX Example 11; Page 28-30; 52pp; English.
XX
XX scFv2(FRP5/225)-ETA (AAW05139) comprises the single-chain binding region
XX of murine monoclonal antibody FRP5 (specific for human epidermal growth
XX factor receptor erbB-2, see also AAW05134) joined to portions of exotoxin
XX A from Pseudomonas aeruginosa and to the single-chain binding region of
XX murine monoclonal antibody 225 (specific for human epidermal growth
XX factor receptor, see also AAW05133). It is encoded by plasmid pMS238-5-
XX 225 (AAT42039). This plasmid can be utilised in the prodn. of the
XX bivalent fusion protein in bacterial (esp. E. coli) host cells. Such
XX fusion proteins (see also AAW05138-44) are useful as antitumour agents
XX
XX Sequence 892 AA:
SQ
Query Match 55.5%; Score 1302; DB 2; Length 892;
Best Local Similarity 99.2%; Pred. No. 2e-82;
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

```

```

AAW05142
ID AAW05142 standard; protein; 895 AA.
XX
XX AAW05142;
AC
XX 29-JAN-1997 (first entry)
XX
XX scFv2(FRP5/FRP5)-ETA (version 1).
XX
XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
XX epidermal growth factor; receptor; plasmid pMS242-5-5; cancer;
XX exotoxin A; ETA; antitumour.
XX
XX Mus; sp.
XX Pseudomonas; aeruginosa.
XX Synthetic.
XX Chimeric.
XX
XX Key
XX Location/Qualifiers
XX 1..21
XX /label= Sig peptide
XX /note= "ompa signal peptide"
XX
XX Peptide
XX 22..44
XX /label= Spacer
XX 45..159
XX /label= ETA
XX /note= "exotoxin A amino acids 252-366"
XX 160..162
XX /label= Spacer
XX 163..402
XX /label= scFv(FRP5)
XX 403..413
XX /label= Spacer
XX 414..644
XX /label= ETA
XX /note= "exotoxin A amino acids 380-610"
XX
XX Peptide
XX 645..646
XX /label= Spacer
XX 647..886
XX /label= scFv(FRP5)
XX 887..895
XX /label= Spacer
XX
XX EP739984-A1.
XX
XX 30-OCT-1996.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX 26-APR-1995; 95EP-00106275.
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX
XX WPI; 1996-478748/48.
XX
XX N-PSDB; AAT42042.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX analogues - and comprise at least two different cell surface binding
XX domain(s), useful for tumour therapy.
XX
XX Example 12; Page 37-39; 52pp; English.
XX
XX scFv2(FRP5/FRP5)-ETA version 1 (AAW05142) includes 2 copies of the single
XX -chain binding region of murine monoclonal antibody FRP5 (specific for
XX human epidermal growth factor receptor erbB-2, see also AAW05134) joined
XX to portions of exotoxin A from Pseudomonas aeruginosa. It is encoded by
XX plasmid pMS242-5-5 (AAT42042). This plasmid can be utilised in the prodn.
XX of the bivalent fusion protein in bacterial (esp. E. coli) host cells.
XX Such fusion proteins (see also AAW05138-44) are useful as antitumour
XX agents
XX

```

SQ Sequence 895 AA:  
 Query Match 55.5%; Score 1302; DB 2; Length 895;  
 Best Local Similarity 99.2%; Pred. No. 2e-82;  
 Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 QIQLVSGPELKKRGETVTKISCKASGYPTNYGMNWKQAPQGLKMWGINTSTGESTF 60  
 DB 163 QVQLQDSGPELKKRGETVTKISCKASGYPTNYGMNWKQAPQGLKMWGINTSTGESTF 222  
 QY 61 ADDPKGRFDFSLFETSAANTAYIQINNLKSEDMATYFCARMEYHGVYPWGQGTIVTVSSG 120  
 DB 223 ADDPKGRFDFSLFETSAANTAYIQINNLKSEDMATYFCARMEYHGVYPWGQGTIVTVSSG 282  
 QY 121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMVYQKPGQSPKL 180  
 DB 283 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMVYQKPGQSPKL 342  
 QY 181 LIYSASSRYTGVPSRFTGSGGSDPFTFTISSVQAEIDLAVYFCQGHFRTPTFGSGTKLEI 240  
 DB 343 LIYSASSRYTGVPSRFTGSGGSDPFTFTISSVQAEIDLAVYFCQGHFRTPTFGSGTKLEI 402  
 QY 241 KALE 244  
 DB 403 KALE 406  
 RESULT 15  
 AAW05144 ID AAW05144 standard; protein; 899 AA.  
 AC AAW05144;  
 DT 29-JAN-1997 (first entry)  
 DE scFv2(FRP5/FRP5)-ETA (version 3).  
 XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;  
 KM epidermal growth factor; receptor; plasmid pMS246-5-5; cancer;  
 KW exotoxin A; ETA; antitumour.  
 XX Mus; gp.  
 OS Pseudomonas; aeruginosa.  
 OS Synthetic.  
 OS Chimeric.  
 XX Key Location/Qualifiers  
 FH Peptide 1..21  
 FT /label= Sig\_peptide  
 FT /note= "ompa signal peptide"  
 FT Peptide 22..38  
 FT /label= Spacer  
 FT Region 39..278  
 FT /label= scFv(FRP5)  
 FT Peptide 279..289  
 FT /label= Spacer  
 FT Region 290..648  
 FT /label= ETA  
 FT /note= "exotoxin A amino acids 252-610"  
 FT Peptide 649..650  
 FT /label= Spacer  
 FT Region 651..890  
 FT /label= scFv(FRP5)  
 FT Peptide 891..899  
 FT /label= Spacer  
 XX EP73984-A1.  
 XX 30-OCT-1996.  
 XX 26-APR-1995;  
 XX 26-APR-1995;  
 PR 95EP-00106275.  
 PR 95EP-00106275.

XX (SANT-) SAN TUMORFORSCHUNGS GMBH.  
 PA Wels W, Schmidt M, Groner B;  
 PI WPI: 1996-478748/48.  
 DR N-PSDB; AAT42044.  
 DR Bivalent fusion proteins that bind epidermal growth factor receptor or  
 PT analogues - and comprise at least two different cell surface binding  
 PT domain(s), useful for tumour therapy.  
 XX Example 12; Page 43-45; 52pp; English.  
 PS scFv2(FRP5/FRP5)-ETA version 3 (AAW05144) includes 2 copies of the single  
 CC chain binding region of murine monoclonal antibody FRP5 (specific for  
 CC human epidermal growth factor receptor erbB-2, see also AAW05134) joined  
 CC to portions of exotoxin A from Pseudomonas aeruginosa. It is encoded by  
 CC plasmid pMS246-5-5 (AAT42044). This plasmid can be utilised in the prodn.  
 CC of the bivalent fusion protein in bacterial (esp. E. coli) host cells.  
 CC Such fusion proteins (see also AAW05138-43) are useful as antitumour  
 CC agents  
 XX Sequence 899 AA:  
 SQ Query Match 55.5%; Score 1302; DB 2; Length 899;  
 Best Local Similarity 99.2%; Pred. No. 2e-82;  
 Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 QIQLVSGPELKKRGETVTKISCKASGYPTNYGMNWKQAPQGLKMWGINTSTGESTF 60  
 DB 39 QVQLQDSGPELKKRGETVTKISCKASGYPTNYGMNWKQAPQGLKMWGINTSTGESTF 98  
 QY 61 ADDPKGRFDFSLFETSAANTAYIQINNLKSEDMATYFCARMEYHGVYPWGQGTIVTVSSG 120  
 DB 99 ADDPKGRFDFSLFETSAANTAYIQINNLKSEDMATYFCARMEYHGVYPWGQGTIVTVSSG 158  
 QY 121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMVYQKPGQSPKL 180  
 DB 159 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMVYQKPGQSPKL 218  
 QY 181 LIYSASSRYTGVPSRFTGSGGSDPFTFTISSVQAEIDLAVYFCQGHFRTPTFGSGTKLEI 240  
 DB 219 LIYSASSRYTGVPSRFTGSGGSDPFTFTISSVQAEIDLAVYFCQGHFRTPTFGSGTKLEI 278  
 QY 241 KALE 244  
 DB 279 KALE 282

Search completed: April 20, 2005, 07:04:53  
 Job time : 66.1759 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 07:01:08 ; Search time 20.0918 Seconds  
(without alignments)  
1645.921 Million cell updates/sec

Title: US-09-596-774-7  
Perfect score: 2345  
Sequence: 1 QIQVQSGPELKKPGETVKI.....LSTATKOTYDALHMOTLAPR 443

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/1/1aa/5A\_COMB.pep:\*  
2: /cgn2\_6/ptodata/1/1aa/5B\_COMB.pep:\*  
3: /cgn2\_6/ptodata/1/1aa/6A\_COMB.pep:\*  
4: /cgn2\_6/ptodata/1/1aa/6B\_COMB.pep:\*  
5: /cgn2\_6/ptodata/1/1aa/PTUS\_COMB.pep:\*  
6: /cgn2\_6/ptodata/1/1aa/backfiles1.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1308	55.8	615	4	US-08-840-713-35
2	1308	55.8	617	4	US-08-840-713-37
3	1303	55.6	711	1	US-08-235-838-7
4	1303	55.6	711	2	US-08-465-473B-7
5	1302	55.5	530	4	US-08-840-713-2
6	1302	55.5	637	1	US-08-235-838-14
7	1302	55.5	637	2	US-08-465-473B-14
8	1284	54.8	241	2	US-08-465-473B-5
9	1284	54.8	250	1	US-08-133-804-2
10	981.5	41.9	250	1	US-08-461-184-8
11	981.5	41.9	250	1	US-08-463-675-8
12	981.5	41.9	250	1	US-08-464-589-8
13	981.5	41.9	250	1	US-08-461-838-2
14	981.5	41.9	250	1	US-08-461-386-2
15	981.5	41.9	250	2	US-08-356-786-16
16	975.5	41.6	622	2	US-08-488-113B-148
17	903.5	38.5	240	1	US-08-477-484B-148
18	903.5	38.5	240	2	US-08-646-360-148
19	903.5	38.5	240	3	US-08-839-765-148
20	903.5	38.5	240	3	US-09-136-389-148
21	903.5	38.5	240	3	US-09-610-838-148
22	903.5	38.5	240	4	US-09-711-485-148
23	903.5	38.5	240	4	US-08-875-811-55
24	891	38.0	366	3	US-08-875-811-55
25	890	38.0	365	3	US-09-485-737B-85
26	889.5	37.9	541	3	US-10-071-485-85
27	889.5	37.9	541	4	US-10-071-485-85

28	888.5	37.9	267	3	US-09-485-737B-2	Sequence 2, Appli
29	888.5	37.9	267	4	US-10-071-485-2	Sequence 2, Appli
30	888.5	37.9	711	3	US-09-485-737B-90	Sequence 90, Appli
31	888.5	37.9	711	4	US-10-071-485-90	Sequence 90, Appli
32	885.5	37.8	259	4	US-09-419-788-115	Sequence 115, App
33	884.5	37.7	284	3	US-09-184-658-40	Sequence 40, Appli
34	884.5	37.7	284	4	US-09-504-262D-40	Sequence 40, Appli
35	878.5	37.5	534	2	US-08-356-786-10	Sequence 10, Appli
36	851	36.3	240	4	US-10-092-246-35	Sequence 35, Appli
37	851	36.3	240	4	US-10-092-246-36	Sequence 36, Appli
38	851	36.3	240	4	US-10-096-246A-35	Sequence 35, Appli
39	851	36.3	240	4	US-10-096-246A-37	Sequence 37, Appli
40	848	36.2	240	3	US-09-485-737B-91	Sequence 91, Appli
41	848	36.2	240	4	US-10-071-485-91	Sequence 91, Appli
42	843	35.9	240	4	US-10-096-246A-36	Sequence 36, Appli
43	841.5	35.9	249	2	US-08-797-689-18	Sequence 18, Appli
44	841.5	35.9	249	4	US-09-984-186-18	Sequence 18, Appli
45	839	35.8	240	4	US-10-092-246-37	Sequence 37, Appli

ALIGNMENTS

```
RESULT 1
US-08-840-713-35
; Sequence 35, Application US/08840713
; Patent No. 6498233
;
; GENERAL INFORMATION:
; APPLICANT: WELS, Winfried, Dr.
; APPLICANT: FOYMINAYA, Jesus
; TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM
; NUMBER OF SEQUENCES: 58
; CORRESPONDENCE ADDRESS:
; ADDRESS: Nikolaio, Marmelstein, Murray & Oram LLP
; STREET: 655 15th St., N.W., Suite 330 - G St. Lobby
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-5701
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/840,713
; FILING DATE: 25-APR-1997
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Kites, Monica Chin
; REGISTRATION NUMBER: 36,105
; REFERENCE/DOCKET NUMBER: 1614-7014
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 638 - 5000
; TELEFAX: (202) 638 - 4810
; INFORMATION FOR SEQ ID NO: 35:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
;
; US-08-840-713-35
;
; Query Match 55.8%; Score 1308; DB 4; Length 615;
; Best Local Similarity 96.5%; Pred. No. 4.3e-100;
; Matches 245; Conservative 3; Mismatches 6; Indels 0; Gaps 0;
;
; QY 1 QIQVQSGPELKKPGETVKISCKASGYPTTYGNWYKQAPGQGLKMMGMINTSGSTF 60
; :|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
; Db 362 QIQVQSGPELKKPGETVKISCKASGYPTTYGNWYKQAPGQGLKMMGMINTSGSTF 421
;
; QY 61 ADDEKGFDFSLSTANTAYLQINNLKSEDMATYFCARWEVYHGVYPMGGITVTYSSG 120
; |||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
```

Db 422 ADDEKGRFDFSLSTANTAYLIQINNLSKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 481  
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVNNAVAMYOQKPGQSPKL 180  
Db 482 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVNNAVAMYOQKPGQSPKL 541  
QY 181 LIYSASRYTGVPSRFTGSGGPDFTFTISSVQAEIDLAVYFCQHFRTPTFGSGTKLEI 240  
Db 542 LIYSASRYTGVPSRFTGSGGPDFTFTISSVQAEIDLAVYFCQHFRTPTFGSGTKLEI 601  
QY 241 KALEISNVMTFSS 254  
Db 602 KALEDLSSERRFSA 615

## RESULT 2

US-08-840-713-37  
Sequence 37, Application US/08840713  
Patent No. 6498233  
GENERAL INFORMATION:  
APPLICANT: WELLS, Winfried, Dr.  
APPLICANT: FOYMINAYA, Jesus  
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby  
CITY: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/840,713  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kiltz, Monica Chin  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: 1614-7014  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638 - 5000  
TELEFAX: (202) 638 - 4810  
INFORMATION FOR SEQ ID NO: 37:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 617 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-840-713-37

Query Match 55.8%; Score 1308; DB 4; Length 617;  
Best Local Similarity 96.5%; Pred. No. 4.3e-100;

Matches 245; Conservative 3; Mismatches 6; Indels 0; Gaps 0;

QY 1 QIQLVSGPELKKPGETVKISCKASGYPTNYGNMVKQAPQGLKMMGMINTSGESTF 60  
Db 364 QVQLQDSGPELKKPGETVKISCKASGYPTNYGNMVKQAPQGLKMMGMINTSGESTF 423  
QY 61 ADDEKGRFDFSLSTANTAYLIQINNLSKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 120  
Db 424 ADDEKGRFDFSLSTANTAYLIQINNLSKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 483  
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVNNAVAMYOQKPGQSPKL 180  
Db 484 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVNNAVAMYOQKPGQSPKL 543  
QY 181 LIYSASRYTGVPSRFTGSGGPDFTFTISSVQAEIDLAVYFCQHFRTPTFGSGTKLEI 240  
Db 542 LIYSASRYTGVPSRFTGSGGPDFTFTISSVQAEIDLAVYFCQHFRTPTFGSGTKLEI 601

Db 544 LIYSASRYTGVPSRFTGSGGPDFTFTISSVQAEIDLAVYFCQHFRTPTFGSGTKLEI 603  
QY 241 KALEISNVMTFSS 254  
Db 604 KALEDLSSERRFSA 617

## RESULT 3

US-08-235-838-7  
Sequence 7, Application US/08235838  
Patent No. 5571894  
GENERAL INFORMATION:  
APPLICANT: Wells, Winfried S.  
APPLICANT: Hynes, Nancy E.  
APPLICANT: Harwerth, Ina-Maria  
APPLICANT: Groner, Bernd  
APPLICANT: Hardman, No. 5571894man  
APPLICANT: Zwickl, Markus  
TITLE OF INVENTION: Recombinant Antibodies Specific for a  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
OPERATING SYSTEM: IBM PC compatible  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,838  
FILING DATE: TBA  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/828,832  
FILING DATE: 31-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 91-810079.3  
FILING DATE: 05-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919)541-8614  
TELEFAX: (919)541-8689  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 711 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
US-08-235-838-7

Query Match 55.6%; Score 1303; DB 1; Length 711;  
Best Local Similarity 94.2%; Pred. No. 1.4e-99;

Matches 245; Conservative 2; Mismatches 1; Indels 12; Gaps 1;

QY 1 QIQLVSGPELKKPGETVKISCKASGYPTNYGNMVKQAPQGLKMMGMINTSGESTF 60  
Db 23 QVQLQDSGPELKKPGETVKISCKASGYPTNYGNMVKQAPQGLKMMGMINTSGESTF 82  
QY 61 ADDEKGRFDFSLSTANTAYLIQINNLSKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 120  
Db 83 ADDEKGRFDFSLSTANTAYLIQINNLSKSEDMATYFCARMEVYHGVPYWGQTTVTVSSG 142  
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVNNAVAMYOQKPGQSPKL 180  
Db 143 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSTTCASQDVNNAVAMYOQKPGQSPKL 202



Qy 181 LIYASASRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTFGSGTKLEI 240  
Db 203 LIYASASRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTFGSGTKLEI 262  
Qy 241 KALEISNSVMYFSSVVPVLO 260  
Db 263 KALE-----PVLE 270

RESULT 4  
US-08-465-473B-7  
; Sequence 7, Application US/08465473B  
; Patent No. 5939531  
; GENERAL INFORMATION:  
; APPLICANT: Wels, Winfried S.  
; APPLICANT: Hynes, Nancy E.  
; APPLICANT: Harwerth, Ina-Maria  
; APPLICANT: Groner, Bernd  
; APPLICANT: Hardman, No. 5939531man  
; APPLICANT: Zwickl, Markus  
; TITLE OF INVENTION: Recombinant Antibodies Specific for a  
; TITLE OF INVENTION: Growth Factor Receptor  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NOVARTIS Corporation  
; STREET: 564 Morris Avenue  
; CITY: Summit  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07901-6940  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,473B  
; FILING DATE: 5 June 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/828,832  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 91-810079.3  
; FILING DATE: 05-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pfeiffer, Heena J.  
; REGISTRATION NUMBER: 22,640  
; REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)522 6940  
; TELEFAX: (908)522 6955  
; INFORMATION FOR SEQ ID NO: 7:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 711 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-465-473B-7

Query Match 55.6%; Score 1303; DB 2; Length 711;  
Best Local Similarity 94.2%; Pred. No. 14e-99;  
Matches 245; Conservative 2; Mismatches 1; Indels 12; Gaps 1;

Qy 1 QIOLVQSGPELKKPGETYKISCKASGYPTNYGNMVAQAPOGGLKMMGMINSTGSESTF 60  
Db 23 QVQLQQSGPELKKPGETYKISCKASGYPTNYGNMVAQAPOGGLKMMGMINSTGSESTF 82  
Qy 61 ADDPKGRDPSLETSANTAYLIQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTVSSG 120  
Db 83 ADDPKGRDPSLETSANTAYLIQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTVSSG 142

Qy 121 GGGSGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAYQKPGQSPKL 180  
Db 143 GGGSGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAYQKPGQSPKL 202  
Qy 181 LIYASASRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTFGSGTKLEI 240  
Db 203 LIYASASRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTFGSGTKLEI 262  
Qy 241 KALEISNSVMYFSSVVPVLO 260  
Db 263 KALE-----PVLE 270

RESULT 5  
US-08-840-713-2  
; Sequence 2, Application US/08840713  
; Patent No. 6498223  
; GENERAL INFORMATION:  
; APPLICANT: WELS, Winfried, Dr.  
; APPLICANT: FOYMINAYA, Jesus  
; TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM  
; NUMBER OF SEQUENCES: 58  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
; STREET: 655 15th St., N.W., Suite 330 - G St. Lobby  
; CITY: Washington  
; STATE: D.C.  
; COUNTRY: USA  
; ZIP: 20005-5701  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/840,713  
; FILING DATE: 25-APR-1997  
; CLASSIFICATION: 514  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Kites, Monica Chiu  
; REGISTRATION NUMBER: 36,105  
; REFERENCE/DOCKET NUMBER: 1614-7014  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (202) 638 - 5000  
; TELEFAX: (202) 638 - 4810  
; INFORMATION FOR SEQ ID NO: 2:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 530 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-840-713-2

Query Match 55.5%; Score 1302; DB 4; Length 530;  
Best Local Similarity 99.2%; Pred. No. 1.1e-99;  
Matches 242; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 QIOLVQSGPELKKPGETYKISCKASGYPTNYGNMVAQAPOGGLKMMGMINSTGSESTF 60  
Db 18 QVQLQQSGPELKKPGETYKISCKASGYPTNYGNMVAQAPOGGLKMMGMINSTGSESTF 77  
Qy 61 ADDPKGRDPSLETSANTAYLIQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTVSSG 120  
Db 78 ADDPKGRDPSLETSANTAYLIQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTVSSG 137  
Qy 121 GGGSGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAYQKPGQSPKL 180  
Db 138 GGGSGGGGGGGGSDIQLTQSHKFLSTSVGDRVSIITCKASQDVYNAVAYQKPGQSPKL 197  
Qy 181 LIYASASRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTFGSGTKLEI 240  
Db 198 LIYASASRYTGVPSRFTSGSGPDFTFTISSVOAEDLAVYFCQOHFRPTFTFGSGTKLEI 257



Db 153 GGGSGGGGGGGSDIQLTOSHKFLSTSVGDRVSITCKASQDVYNAVAYOQKPGQSPKL 212  
Qy 181 LIYASSRVTGVPSPRFTGSGGPDFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 240  
Db 213 LIYASSRVTGVPSPRFTGSGGPDFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 272  
Qy 241 KALE 244  
273 KALE 276

## RESULT 8

US-08-235-838-5  
; Sequence 5, Application US/08235838  
; Patent No. 5571894  
; GENERAL INFORMATION:  
; APPLICANT: Weis, Winfried S.  
; APPLICANT: Hynes, Nancy E.  
; APPLICANT: Harwerth, Ina-Maria  
; APPLICANT: Groner, Bernd  
; APPLICANT: Hardman, No. 5571894man  
; APPLICANT: Zwickl, Markus  
; TITLE OF INVENTION: Recombinant Antibodies Specific for a  
; NUMBER OF SEQUENCES: 16  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: CIBA-GEIGY Corporation  
; STREET: 7 Skyline Drive  
; CITY: Hawthorne  
; STATE: New York  
; COUNTRY: USA  
; ZIP: 10532  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/235,838  
; FILING DATE: TBA  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/828,832  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 91-810079.3  
; FILING DATE: 05-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Elmer, James Scott  
; REGISTRATION NUMBER: 36,129  
; REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (919)541-8614  
; TELEFAX: (919)541-8689  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 241 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-235-838-5

Query Match 54.8%; Score 1284; DB 1; Length 241;  
Best Local Similarity 99.2%; Pred. No. 1.2e-98;  
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 QIOLVSGPELKKRGETIVKISCKASGYPTNYGNMNVKQAPGQGLKMMGMINSTGSESTF 60  
Db 2 QVQLQDSGPPELKKRGETIVKISCKASGYPTNYGNMNVKQAPGQGLKMMGMINSTGSESTF 61  
Qy 61 ADDEKGRFDSLETSANTAVYLIQINNLSKSDMATYFCARMEVYHGVPYWGQGTITVTVSSG 120  
Db 62 ADDEKGRFDSLETSANTAVYLIQINNLSKSDMATYFCARMEVYHGVPYWGQGTITVTVSSG 121

Qy 121 GGGSGGGGGGGSDIQLTOSHKFLSTSVGDRVSITCKASQDVYNAVAYOQKPGQSPKL 180  
Db 122 GGGSGGGGGGGSDIQLTOSHKFLSTSVGDRVSITCKASQDVYNAVAYOQKPGQSPKL 181  
Qy 181 LIYASSRVTGVPSPRFTGSGGPDFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 240  
Db 182 LIYASSRVTGVPSPRFTGSGGPDFTTISVQAEADLAVYFCQGHFRTPFTFGSGTKLEI 241

## RESULT 9

US-08-465-473B-5  
; Sequence 5, Application US/08465473B  
; Patent No. 5939531  
; GENERAL INFORMATION:  
; APPLICANT: Weis, Winfried S.  
; APPLICANT: Hynes, Nancy E.  
; APPLICANT: Harwerth, Ina-Maria  
; APPLICANT: Groner, Bernd  
; APPLICANT: Hardman, No. 5939531man  
; APPLICANT: Zwickl, Markus  
; TITLE OF INVENTION: Recombinant Antibodies Specific for a  
; NUMBER OF SEQUENCES: 34  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: NOVARTIS Corporation  
; STREET: 564 Morris Avenue  
; CITY: Summit  
; STATE: New Jersey  
; COUNTRY: USA  
; ZIP: 07901-6940  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patentin Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/465,473B  
; FILING DATE: 5 June 1995  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 07/828,832  
; FILING DATE: 31-JAN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: GB 91-810079.3  
; FILING DATE: 05-FEB-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Pfeiffer, Heena J.  
; REGISTRATION NUMBER: 22,640  
; REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (908)522 6940  
; TELEFAX: (908)522 6955  
; INFORMATION FOR SEQ ID NO: 5:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 241 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; US-08-465-473B-5

Query Match 54.8%; Score 1284; DB 2; Length 241;  
Best Local Similarity 99.2%; Pred. No. 1.2e-98;  
Matches 238; Conservative 1; Mismatches 1; Indels 0; Gaps 0;  
Qy 1 QIOLVSGPELKKRGETIVKISCKASGYPTNYGNMNVKQAPGQGLKMMGMINSTGSESTF 60  
Db 2 QVQLQDSGPPELKKRGETIVKISCKASGYPTNYGNMNVKQAPGQGLKMMGMINSTGSESTF 61  
Qy 61 ADDEKGRFDSLETSANTAVYLIQINNLSKSDMATYFCARMEVYHGVPYWGQGTITVTVSSG 120  
Db 62 ADDEKGRFDSLETSANTAVYLIQINNLSKSDMATYFCARMEVYHGVPYWGQGTITVTVSSG 121



RESULT 12

US-08-463-675-8

Sequence 8, Application US/08463675

Patent No. 5658763

GENERAL INFORMATION:

APPLICANT: DORAI, HAIMANTI

APPLICANT: OPPERMANN, HERMANN

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN

TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC

STREET: 45 SOUTH STREET

CITY: HOPKINTON

STATE: MA

COUNTRY: USA

ZIP: 07148

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/463.675

FILING DATE: 05-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: 08/143,498

FILING DATE: 25-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: KELLEY, ROBIN D

REGISTRATION NUMBER: 34,637

REFERENCE/DOCKET NUMBER: CRP093

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/248-7000

TELEFAX: 617/248-7100

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 250 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-463-675-8

Query Match 41.9%; Score 981.5; DB 1; Length 250;

Best Local Similarity 76.8%; Pred. No. 1.6e-73;

Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 1 QIQVOSGPELKKKGEYTKISCKASGYPTNYGMMWVQAQPGQGLKMMGMINSTGSESTF 60

DB 3 EIQLVOSGPELKKKGEYTKISCKASGYPTNYGMMWVQAQPGQGLKMMGMINSTGSESTF 62

QY 61 ADFKGRFDFSLFSTANTAYLQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTSSG 120

DB 63 AEFKGRFDFSLFSTANTAYLQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTSSG 121

QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVYNAVAMYOQKPGQSPKL 180

DB 122 SSSSGSSSSGSSSDIQLTQSHKFLSTSVGRVSIITCKASQDVYNAVAMYOQKPGQSPKL 181

QY 181 LIYASASRYTGVPRFTSGSGSDPTFTTSSVQAEADLAVYFCQOHFRPPTFGSGTKLEI 240

DB 182 LIYWTSTRHTGVPRFTSGSGSDPTFTTSSVQAEADLAVYFCQOHFRPPTFGSGTKLEI 241

QY 241 K 241

DB 242 K 242

RESULT 13

US-08-464-589-8

Sequence 8, Application US/08464589

Patent No. 5733782

GENERAL INFORMATION:

APPLICANT: DORAI, HAIMANTI

APPLICANT: OPPERMANN, HERMANN

TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN

TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA

NUMBER OF SEQUENCES: 10

CORRESPONDENCE ADDRESS:

ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC

STREET: 45 SOUTH STREET

CITY: HOPKINTON

STATE: MA

COUNTRY: USA

ZIP: 07148

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/464,589

FILING DATE: 05-JUN-1995

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US 08/143,498

FILING DATE: 25-OCT-1993

ATTORNEY/AGENT INFORMATION:

NAME: KELLEY, ROBIN D

REGISTRATION NUMBER: 34,637

REFERENCE/DOCKET NUMBER: CRP093

TELECOMMUNICATION INFORMATION:

TELEPHONE: 617/248-7000

TELEFAX: 617/248-7100

INFORMATION FOR SEQ ID NO: 8:

SEQUENCE CHARACTERISTICS:

LENGTH: 250 amino acids

TYPE: amino acid

TOPOLOGY: linear

MOLECULE TYPE: protein

US-08-464-589-8

Query Match 41.9%; Score 981.5; DB 1; Length 250;

Best Local Similarity 76.8%; Pred. No. 1.6e-73;

Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 1 QIQVOSGPELKKKGEYTKISCKASGYPTNYGMMWVQAQPGQGLKMMGMINSTGSESTF 60

DB 3 EIQLVOSGPELKKKGEYTKISCKASGYPTNYGMMWVQAQPGQGLKMMGMINSTGSESTF 62

QY 61 ADFKGRFDFSLFSTANTAYLQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTSSG 120

DB 63 AEFKGRFDFSLFSTANTAYLQINNLSKEDMATYFCARMEVYHGVVPYWGQTTVTSSG 121

QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVYNAVAMYOQKPGQSPKL 180

DB 122 SSSSGSSSSGSSSDIQLTQSHKFLSTSVGRVSIITCKASQDVYNAVAMYOQKPGQSPKL 181

QY 181 LIYASASRYTGVPRFTSGSGSDPTFTTSSVQAEADLAVYFCQOHFRPPTFGSGTKLEI 240

DB 182 LIYWTSTRHTGVPRFTSGSGSDPTFTTSSVQAEADLAVYFCQOHFRPPTFGSGTKLEI 241

QY 241 K 241

DB 242 K 242

RESULT 14

US-08-461-838-2

Sequence 2, Application US/08461838

Patent No. 5753204

GENERAL INFORMATION:

APPLICANT: HUSTON, JAMES S.

APPLICANT: OPPERMANN, HERMANN

APPLICANT: HUSTON, L. L.

```

; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,838
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-838-2

Query Match      41.9%; Score 981.5; DB 1; Length 250;
Best Local Similarity 76.8%; Pred. No. 1.6e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 1 QIOLVSGPELKKPGETVYKISCKASGYPTNYGMNVKQAPGQGLKMWGMINTSGSTF 60
Db 3 EIQLVSGPELKKPGETVYKISCKASGYPTNYGMNVKQAPGQGLKMWGMINTSGSTF 62
QY 61 ADFPKGRFDSLETSANTAYLQINNLKSEDMATYFCARWEVYHGVPYVGQGTIVVSSG 120
Db 63 AEFKGRFAFSLETSASTAYLQINNLKSEDMATYFCGRQFTYGGFANWGQGLTVVSA- 121
QY 121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSITCKASODVYNNAVYQOKPGQSPKL 180
Db 122 SSSSGSSSSGSSSDIVMTQSPKFMSTSVGDRVSISCKASQDVSTAVAMVQOKPGQSPKL 181
QY 181 LIYSASRRTGVPSRFTGSGSPDFTTISVQAEIDLAVYFCQGHFRPTFTGSGTKLEI 240
Db 182 LIYWTSTRHTGVPRFTGSGSGTDYTLTISVQAEIDLALHYCOQHYRVPYTFGGGTLEI 241
QY 241 K 241
Db 242 K 242

RESULT 15
US-08-461-386-2
; Sequence 2, Application US/08461386
; Patent No. 5837846
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
```

```

; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: IBM PC compatible
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,386
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 250 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-461-386-2

Query Match      41.9%; Score 981.5; DB 2; Length 250;
Best Local Similarity 76.8%; Pred. No. 1.6e-73;
Matches 185; Conservative 22; Mismatches 33; Indels 1; Gaps 1;

QY 1 QIOLVSGPELKKPGETVYKISCKASGYPTNYGMNVKQAPGQGLKMWGMINTSGSTF 60
Db 3 EIQLVSGPELKKPGETVYKISCKASGYPTNYGMNVKQAPGQGLKMWGMINTSGSTF 62
QY 61 ADFPKGRFDSLETSANTAYLQINNLKSEDMATYFCARWEVYHGVPYVGQGTIVVSSG 120
Db 63 AEFKGRFAFSLETSASTAYLQINNLKSEDMATYFCGRQFTYGGFANWGQGLTVVSA- 121
QY 121 GGGSGGGSGGGGSDIQLTQSHKFLSTSVGDRVSITCKASODVYNNAVYQOKPGQSPKL 180
Db 122 SSSSGSSSSGSSSDIVMTQSPKFMSTSVGDRVSISCKASQDVSTAVAMVQOKPGQSPKL 181
QY 181 LIYSASRRTGVPSRFTGSGSPDFTTISVQAEIDLAVYFCQGHFRPTFTGSGTKLEI 240
Db 182 LIYWTSTRHTGVPRFTGSGSGTDYTLTISVQAEIDLALHYCOQHYRVPYTFGGGTLEI 241
QY 241 K 241
Db 242 K 242
```

Search completed: April 20, 2005, 07:08:36  
Job time : 21.0918 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 07:02:38 ; Search time 56.8451 Seconds  
(without alignments)  
2590.199 Million cell updates/sec

Title: US-09-596-774-7  
Perfect score: 2345  
Sequence: 1 QIOLVSGPELKKPGETVKI.....LSTATKDYALHMQLAPR 443

Scoring table: BLOSUM62  
Gapop 10.0, Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : Published Applications NA:  
1: /cgn2\_6/ptodata/1/pubppaa/US07\_PUBCOMB.pep:\*  
2: /cgn2\_6/ptodata/1/pubppaa/PCT\_NEW\_PUB.pep:\*  
3: /cgn2\_6/ptodata/1/pubppaa/US06\_NEW\_PUB.pep:\*  
4: /cgn2\_6/ptodata/1/pubppaa/US06\_PUBCOMB.pep:\*  
5: /cgn2\_6/ptodata/1/pubppaa/US07\_NEW\_PUB.pep:\*  
6: /cgn2\_6/ptodata/1/pubppaa/PCTUS\_PUBCOMB.pep:\*  
7: /cgn2\_6/ptodata/1/pubppaa/US08\_NEW\_PUB.pep:\*  
8: /cgn2\_6/ptodata/1/pubppaa/US09\_PUBCOMB.pep:\*  
9: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep:\*  
10: /cgn2\_6/ptodata/1/pubppaa/US09B\_PUBCOMB.pep:\*  
11: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*  
12: /cgn2\_6/ptodata/1/pubppaa/US09C\_PUBCOMB.pep:\*  
13: /cgn2\_6/ptodata/1/pubppaa/US10\_PUBCOMB.pep:\*  
14: /cgn2\_6/ptodata/1/pubppaa/US10B\_PUBCOMB.pep:\*  
15: /cgn2\_6/ptodata/1/pubppaa/US10C\_PUBCOMB.pep:\*  
16: /cgn2\_6/ptodata/1/pubppaa/US10D\_PUBCOMB.pep:\*  
17: /cgn2\_6/ptodata/1/pubppaa/US10\_NEW\_PUB.pep:\*  
18: /cgn2\_6/ptodata/1/pubppaa/US11\_NEW\_PUB.pep:\*  
19: /cgn2\_6/ptodata/1/pubppaa/US60\_NEW\_PUB.pep:\*  
20: /cgn2\_6/ptodata/1/pubppaa/US60\_PUBCOMB.pep:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match Length	ID	Description
1	1249	53.3	631	15 US-10-120-198B-2
2	1015.5	43.3	443	13 US-10-006-773-2
3	1015.5	43.3	443	13 US-10-006-773A-2
4	961.5	41.9	250	9 US-09-887-853-2
5	969.5	41.3	250	17 US-10-683-547-2
6	930	39.7	444	8 US-08-812-393A-2
7	928	39.6	449	10 US-09-774-681-2
8	917.5	39.1	505	15 US-10-239-656-79
9	913	38.9	267	9 US-09-766-543-10
10	903.5	38.5	240	14 US-10-127-890-148
11	903.5	38.5	240	14 US-10-127-890-148
12	897	38.3	276	9 US-09-766-543-12
13	889.5	37.9	541	14 US-10-071-485-85

14	888.5	37.9	267	14	US-10-071-485-2	Sequence 2, Appl1
15	888.5	37.9	701	14	US-10-071-485-90	Sequence 90, Appl1
16	886.5	37.8	511	15	US-10-239-656-75	Sequence 79, Appl1
17	875.5	37.3	331	14	US-10-059-261-169	Sequence 169, App
18	872.5	37.2	243	17	US-10-879-994-10	Sequence 10, Appl1
19	872.5	37.2	243	17	US-10-610-452-10	Sequence 2, Appl1
20	859	36.6	622	15	US-10-378-832A-2	Sequence 1104, Ap
21	857	36.5	248	10	US-09-880-748-1104	Sequence 1104, Ap
22	857	36.5	248	15	US-10-293-418-1104	Sequence 1921, Ap
23	854.5	36.4	251	10	US-09-880-748-1921	Sequence 1921, Ap
24	854.5	36.4	251	15	US-10-293-418-1921	Sequence 1921, Ap
25	854.5	36.4	503	14	US-10-239-656-77	Sequence 77, Appl1
26	851	36.3	240	15	US-10-098-246-35	Sequence 35, Appl1
27	851	36.3	240	14	US-10-096-246-36	Sequence 36, Appl1
28	849.5	36.2	249	15	US-10-239-656-73	Sequence 91, Appl1
29	848	36.2	240	14	US-10-071-485-91	Sequence 91, Appl1
30	848	36.2	250	10	US-09-880-748-932	Sequence 932, App
31	846.5	36.1	249	15	US-10-293-418-932	Sequence 932, App
32	846.5	36.1	249	10	US-09-880-748-926	Sequence 926, App
33	846.5	36.1	249	15	US-10-293-418-926	Sequence 926, App
34	846	36.1	248	10	US-09-880-748-1446	Sequence 1446, Ap
35	846	36.1	248	15	US-10-293-418-1446	Sequence 1446, Ap
36	845	36.0	248	10	US-09-880-748-1008	Sequence 1008, Ap
37	845	36.0	248	15	US-09-880-748-1008	Sequence 1008, Ap
38	845	36.0	248	10	US-10-293-418-1008	Sequence 1008, Ap
39	845	36.0	248	15	US-10-293-418-1778	Sequence 1778, Ap
40	842.5	35.9	249	10	US-09-880-748-918	Sequence 918, App
41	842.5	35.9	249	15	US-10-293-418-918	Sequence 918, App
42	841.5	35.9	249	9	US-09-984-186-18	Sequence 18, Appl1
43	841.5	35.9	249	14	US-10-237-708-18	Sequence 18, Appl1
44	841.5	35.9	249	14	US-10-237-708-18	Sequence 18, Appl1
45	841.5	35.9	249	14	US-10-237-866-18	Sequence 18, Appl1

ALIGNMENTS

RESULT 1  
US-10-120-198B-2  
Sequence 2, Application US/10120198B  
Publication No. US20030215427A1  
GENERAL INFORMATION:  
APPLICANT: Jensen, Michael  
TITLE OF INVENTION: CE7-SPECIFIC REDIRECTED IMMUNE CELLS  
FILE REFERENCE: 1954-337  
CURRENT APPLICATION NUMBER: US/10/120.198B  
CURRENT FILING DATE: 2002-04-11  
PRIOR APPLICATION NUMBER: 60/282,859  
PRIOR FILING DATE: 2001-04-11  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 2  
LENGTH: 631  
TYPE: PRT  
ORGANISM: artificial sequence  
FEATURE:  
OTHER INFORMATION: mouse-human chimera  
US-10-120-198B-2

Query Match 53.3%; Score 1249; DB 15; Length 631;  
Best Local Similarity 44.1%; Pred. No. 1,3e-79;  
Matches 274; Conservative 60; Mismatches 96; Indels 192; Gaps 13;  
QY 1 QIOLVSGPELKKPGETVKISCAASGYPTNNYGNWYKQAPGGLKMKMGWINTSGSTF 60  
23 QVLOQPGALVYKGVKSCASGYTFTGWNHMKRPGHGLEWIGETINPSNGRTNY 82  
QY 61 ADPFKGFDSLSNTAYLQINNLSKEDMAYFYFCARWVYHG---YVYWGQTTVTY 117  
83 NERKSAATLTVDSSSTTAPMQLSGLTSBDAVYFCAR--DYTGTSYNPDYWGQTLTV 140  
QY 118 SSGGSGGSGGSGGSGGSDIOLTQSHKFLSTSVGDRVSYTCASQDVYNAVWYQKPGOS 177

```

Db 141 SSGGGSGGGSGGGSDIOMTQSSSFSVSLGDRVITITCKANEDINNRLAMYOQTGNS 200
Qy 178 PKLLIYASAKRYTGVPSRFTGSGSPDFTTITSSQADDLAVYFCQGFRTPTFTGSGTK 237
Db 201 PRLISGATNLVTVGVSFRSGSGSKDYTLITSLQADDFATVYCOQYWSPTFTGSGTE 260
Qy 238 LEIK-----ALEISNVWFSVVPVLQ-- 260
Db 261 LEIKVEKPSDKHTHCPCRPABELLGSPVLFPPKPDITMISTKTPBEVTCVVDVSHED 320
Qy 261 -----KVNSTTKP-----VLR-----TPS 275
Db 321 PEVKNWYVDGVEVNAKTKREEQYNSTYRVSVLTVLHODWLNKEKCKSVSKALPA 380
Qy 276 PVHPT---GTSOPQRP-----DCRPGS 296
Db 381 PIEKTISAKQGPPEPVYTLPPSRDELTKQVSLTCLVKGFPSDIAVEMESNGQPENN 440
Qy 297 VKGTG--LDPLEDPKLCYLDLILFY----- 321
Db 441 YKTPPVLD-----SDGSFELYSKLTVDSKRMQGNVFGSCVWHEALHNHTYQS 490
Qy 322 -----GVII--TALYLAKFSRSATAANLDPNQLYNELNIGRRE 361
Db 491 LSLSPGMALIVLGVAIGLLFIGIFFRVYKFSRSADAPAYQOGQNLVNEINIGRRE 550
Qy 362 YDVEKKRARDPEMGKQRRRNPOEGVYNALQDKMAEYSEIGTGERRGKHGDL 421
Db 551 YDVLDRKRGRPEMGK-PRKNPOEGLYNELQDKMAEYSEIGMGERRGKHGDL 609
Qy 422 QGLSTATDYDALHMOTLAPR 443
Db 610 QGLSTATDYDALHMOTLAPR 631

RESULT 2
US-10-006-773-2
; Sequence 2, Application US/10006773
; Publication No. US20020132983A1
; GENERAL INFORMATION:
; APPLICANT: Jungmans, Richard P.
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti-
; FILE REFERENCE: 003
; CURRENT APPLICATION NUMBER: US/10/006, 773
; CURRENT FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: 60/250, 089
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 19
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens and Mus sp.
US-10-006-773-2

```

```

Query Match 43.3%; Score 1015.5; DB 13; Length 443;
Best Local Similarity 52.0%; Pred. No. 2,4e-63;
Matches 225; Conservative 25; Mismatches 58; Indels 125; Gaps 8;

```

```

Qy 131 GGGSDIQLTQSHKFLTSVGDVRSITCKASQDVYNAVAMVYQKRGSPKLLIYSASRYT 190
Db 16 GVHSDIQLTQSPSSLSASVGDVRYITCKASQDVGVSVAMVYQKRGKAPKLLIYWTSTRT 75
Qy 191 GVPSRFTGSGGPDFTFTTISVQAEIDLAVYFCQGH--FRTPTFGSGTKLEIK----- 241
Db 76 GVPSRFTGSGGPDFTFTTISLQPEDIAITYCOQYSLYR--SFGQGTVEIKRGSGSG 132
Qy 242 -----ALEISNV----- 249
Db 133 GSGSGSGSEVOLVESGGVVGPERSLRLSCSASGPDFTTYMWSWVQAPGKLEWIGE 192
Qy 250 -----MYFS 254

```

```

Db 193 HPDSSTINAPSLKDRFTISRDNAKNTLFLQMSLRPEDTGVYFCASLYRGPFWAYMG 252
Qy 255 VVPLQKVNSTTKKVLRTSPVHPTGTSOP--QRPEDCPR--GSYKGTGLDPLEDPK 310
Db 253 GTPVTVSAPKPTTPAPRPPTPA-PTIASQPLSLRPEARPAAGAVHTRGDLPALDPK 311
Qy 311 CYLLDGLIFTYGVITLALYRAKFSRSATAANLQDPNQLYNELNIGRREYVLEKRA 370
Db 312 CYLLDGLIFTYGVITLALFLVRFKFSRABPPAYQOGQNLVNEINIGRREYVLDKRG 371
Qy 371 RDEPMGKQRRRNPOEGVYNALQDKMAEYSEIGTGERRGKHGDLGYOGISTATKD 430
Db 372 RDEPMGK-PRKNPOGLVNEINQDKMAEYSEIGMGERRGKHGDLGYOGISTATKD 430
Qy 431 TYDALHMOTLAPR 443
Db 431 TYDALHMOTLAPR 443

```

```

RESULT 3
US-10-006-771A-2
; Sequence 2, Application US/10006771A
; Publication No. US20020165360A1
; GENERAL INFORMATION:
; APPLICANT: Jungmans, Richard P.
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen
; FILE REFERENCE: 002
; CURRENT APPLICATION NUMBER: US/10/006, 771A
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: 60/250, 090
; PRIOR FILING DATE: 2000-11-30
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 443
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-006-771A-2

```

```

Query Match 43.3%; Score 1015.5; DB 13; Length 443;
Best Local Similarity 52.0%; Pred. No. 2,4e-63;
Matches 225; Conservative 25; Mismatches 58; Indels 125; Gaps 8;

```

```

Qy 131 GGGSDIQLTQSHKFLTSVGDVRSITCKASQDVYNAVAMVYQKRGSPKLLIYSASRYT 190
Db 16 GVHSDIQLTQSPSSLSASVGDVRYITCKASQDVGVSVAMVYQKRGKAPKLLIYWTSTRT 75
Qy 191 GVPSRFTGSGGPDFTFTTISVQAEIDLAVYFCQGH--FRTPTFGSGTKLEIK----- 241
Db 76 GVPSRFTGSGGPDFTFTTISLQPEDIAITYCOQYSLYR--SFGQGTVEIKRGSGSG 132
Qy 242 -----ALEISNV----- 249
Db 133 GSGSGSGSEVOLVESGGVVGPERSLRLSCSASGPDFTTYMWSWVQAPGKLEWIGE 192
Qy 250 -----MYFS 254
Db 193 HPDSSTINAPSLKDRFTISRDNAKNTLFLQMSLRPEDTGVYFCASLYRGPFWAYMG 252
Qy 255 VVPLQKVNSTTKKVLRTSPVHPTGTSOP--QRPEDCPR--GSYKGTGLDPLEDPK 310
Db 253 GTPVTVSAPKPTTPAPRPPTPA-PTIASQPLSLRPEARPAAGAVHTRGDLPALDPK 311
Qy 311 CYLLDGLIFTYGVITLALYRAKFSRSATAANLQDPNQLYNELNIGRREYVLEKRA 370
Db 312 CYLLDGLIFTYGVITLALFLVRFKFSRABPPAYQOGQNLVNEINIGRREYVLDKRG 371
Qy 371 RDEPMGKQRRRNPOEGVYNALQDKMAEYSEIGTGERRGKHGDLGYOGISTATKD 430
Db 372 RDEPMGK-PRKNPOEGLYNELQDKMAEYSEIGMGERRGKHGDLGYOGISTATKD 430
Qy 431 TYDALHMOTLAPR 443

```





MEDIUM TYPE: Diskette  
 COMPUTER: IBM Compatible  
 OPERATING SYSTEM: DOS  
 SOFTWARE: FastSeq for Windows Version 2.0  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/812.393A  
 FILING DATE: 05-MAR-1997  
 CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER:  
 FILING DATE:  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Murashige, Kate H  
 REGISTRATION NUMBER: 29,959  
 REFERENCE/DOCKET NUMBER: 31333-20001.00  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 202-887-1500  
 TELEFAX: 202-822-0168  
 TELEX:  
 INFORMATION FOR SEQ ID NO: 2:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 444 amino acids  
 TYPE: amino acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: protein  
 FRAGMENT TYPE: internal  
 US-08-812-393A-2

Query Match 39.7%; Score 930; DB 8; Length 444;  
 Best Local Similarity 48.2%; Pred. No. 2.6e-57;  
 Matches 217; Conservative 50; Mismatches 145; Indels 38; Gaps 14;

QY 3 QLVOSGPE--LKKRGETVTKISCKASGVPTNYGMNVKQAPGGLKMMGWINTSGESTF 60  
 DB 24 QVQVQSPLSLVLOEGENALQCSFS--IFTNQ-VQMFYQRPGRSLVSLNPSGTQS-- 78  
 QY 61 ADDEKGRFDFSLTSANTAYLQINNLSKSEDMATYFCARMEVYHGVYVPGQGTVTYSSG 120  
 DB 79 -----GLTSTVTYIKERRSLHSSQITDSGTYLCASNSGGSNAKLTFGKGTLSVKS 133  
 QY 121 GGGSGGGSGGGSGSDIQLTQSHKFLSTVGDRVSTTCASQDVYNAVAYQKFGQSPKL 180  
 DB 134 GGGSGGGSGGGSGSEAVTQSPRNKVAVTGKVTLSQNTNN--HNNMYWYRQDTGHGLRL 192  
 QY 181 LIYS---ASSRYTVGVSPRFTGS--GSGPDFFTTSSVQAEIDLAVYFC---QQHRTPTFTG 233  
 DB 193 IHSYAGSTEKGIIDGYKASRPSQENFSLILELATPSQTSVYFCASGETGTNERLFP 252  
 QY 234 SGTLEIKALEISNSVMYFSSVVPVLQKVNSTTKPVLRTPSPVHPTGTSQORPEDCRP 293  
 DB 253 HGTGLSVLT---SNSIMYFHFVPLPAKPTTT--PAPRPPTPA--PTIASQPL---SLRP 304  
 QY 294 RGSVKGTDLPLEPRKCYLLDGLFTYGVITLALYLRKAFSRSATANLQDPNQLYNE 353  
 DB 305 SSS-----RDPKLCYLLDGLFTYGVITLALYLRKAFSRSADAPAYQOQGNQLYNE 355  
 QY 354 LNLGRREYDVLLEKRAPDEMGKQOARRNPQEGVYNALQDKMAEAYSEIGTKGERRR 413  
 DB 356 LNLGRREYDVLQKRRGRPEMGK--PRKKNPQEGLYNELQDKMAEAYSEIGTKGERRR 414  
 QY 414 GKGHDLGYGLSTAYTKDTYDALHMOTLAPR 443  
 DB 415 GKGHDLGYGLSTAYTKDTYDALHMOTLAPR 444

RESULT 7  
 US-09-774-681-2  
 ; Sequence 2, Application US/09774681  
 ; Publication No. US20030208780A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Sunol Molecular Corporation  
 ; APPLICANT: Sherman, Linda

APPLICANT: Lustgarten, Joseph  
 TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL  
 TITLE OR INVENTION: RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR ANTIGENS  
 FILE REFERENCE: 31333-20001.01  
 CURRENT APPLICATION NUMBER: US/09/774,681  
 CURRENT FILING DATE: 2001-02-01  
 PRIOR APPLICATION NUMBER: US 08/812,393  
 PRIOR FILING DATE: 1997-03-05  
 PRIOR APPLICATION NUMBER: US 60/012,845  
 PRIOR FILING DATE: 1996-03-05  
 NUMBER OF SEQ ID NOS: 65  
 SOFTWARE: FastSeq for Windows Version 4.0  
 SEQ ID NO 2  
 LENGTH: 449  
 TYPE: PRT  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Deduced amino acid derivative of effective T cell  
 ; OTHER INFORMATION: receptor  
 US-09-774-681-2

Query Match 39.6%; Score 928; DB 10; Length 449;  
 Best Local Similarity 48.2%; Pred. No. 3.6e-57;  
 Matches 217; Conservative 50; Mismatches 145; Indels 38; Gaps 14;

QY 3 QLVOSGPE--LKKRGETVTKISCKASGVPTNYGMNVKQAPGGLKMMGWINTSGESTF 60  
 DB 24 QVQVQSPLSLVLOEGENALQCSFS--IFTNQ-VQMFYQRPGRSLVSLNPSGTQS-- 78  
 QY 61 ADDEKGRFDFSLTSANTAYLQINNLSKSEDMATYFCARMEVYHGVYVPGQGTVTYSSG 120  
 DB 79 -----GLTSTVTYIKERRSLHSSQITDSGTYLCASNSGGSNAKLTFGKGTLSVKS 133  
 QY 121 GGGSGGGSGGGSGSDIQLTQSHKFLSTVGDRVSTTCASQDVYNAVAYQKFGQSPKL 180  
 DB 134 GGGSGGGSGGGSGSEAVTQSPRNKVAVTGKVTLSQNTNN--HNNMYWYRQDTGHGLRL 192  
 QY 181 LIYS---ASSRYTVGVSPRFTGS--GSGPDFFTTSSVQAEIDLAVYFC---QQHRTPTFTG 233  
 DB 193 IHSYAGSTEKGIIDGYKASRPSQENFSLILELATPSQTSVYFCASGETGTNERLFP 252  
 QY 234 SGTLEIKALEISNSVMYFSSVVPVLQKVNSTTKPVLRTPSPVHPTGTSQORPEDCRP 293  
 DB 253 HGTGLSVLT---SNSIMYFHFVPLPAKPTTT--PAPRPPTPA--PTIASQPL---SLRP 304  
 QY 294 RGSVKGTDLPLEPRKCYLLDGLFTYGVITLALYLRKAFSRSATANLQDPNQLYNE 353  
 DB 305 SSS-----RDPKLCYLLDGLFTYGVITLALYLRKAFSRSADAPAYQOQGNQLYNE 355  
 QY 354 LNLGRREYDVLLEKRAPDEMGKQOARRNPQEGVYNALQDKMAEAYSEIGTKGERRR 413  
 DB 356 LNLGRREYDVLQKRRGRPEMGK--PRKKNPQEGLYNELQDKMAEAYSEIGTKGERRR 414  
 QY 414 GKGHDLGYGLSTAYTKDTYDALHMOTLAPR 443  
 DB 415 GKGHDLGYGLSTAYTKDTYDALHMOTLAPR 444

RESULT 8  
 US-10-239-656-79  
 ; Sequence 79, Application US/10239656  
 ; Publication No. US20040038339A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KUPER, PETER  
 ; APPLICANT: RIETMULLER, GERT  
 ; APPLICANT: LUTTERBOUSE, RALF  
 ; APPLICANT: BORSCHERT, KATRIN  
 ; APPLICANT: KISCHEL, ROMAN  
 ; APPLICANT: MAYER, MONIKA  
 ; APPLICANT: HOMERISTER, ROBERT  
 TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE  
 ; TITLE OF INVENTION: TO AN EPITOPE OF THE NKGD2 RECEPTOR COMPLEX  
 ; FILE REFERENCE: 029976/0106

OTHER INFORMATION: IL-2-antibody fusions  
US-09-766-543-10  
Query Match 38.9%; Score 913; DB 9; Length 267;  
Best Local Similarity 68.4%; Pred. No. 2.3e-56;  
Matches 167; Conservative 37; Mismatches 36; Indels 4; Gaps 2;

CURRENT APPLICATION NUMBER: US/10/239,656  
CURRENT FILING DATE: 2003-03-06  
PRIOR APPLICATION NUMBER: PCT/EP01/03414  
PRIOR FILING DATE: 2001-03-26  
PRIOR APPLICATION NUMBER: EP 00106467.4  
PRIOR FILING DATE: 2000-03-24  
NUMBER OF SEQ ID NOS: 92  
SOFTWARE: PatentIn Ver. 2.1  
SEQ ID NO 79  
LENGTH: 505  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: Synthetic 3810xps-  
US-10-239-656-79

Query Match 39.1%; Score 917.5; DB 15; Length 505;  
Best Local Similarity 50.5%; Pred. No. 2.2e-56;  
Matches 185; Conservative 25; Mismatches 31; Indels 125; Gaps 4;

QY 1 QIQVQSGPELKKRGETVTKISCKASGYPTNYGMNVRQAQPGQGLKMWGINTSTGESE 60  
DB 128 QVQLQSGPELKKRGETVTKISCKASGYPTNYGMNVRQAQPGQGLKMWGINTSTGESE 187  
QY 61 ADPFKGRPFSLTSANTAYLQINNLSKSEDMATYFCAR----- 98  
DB 188 GDDPKGRPFSLTSANTAYLQINNLSKSEDMATYFCARFTSPDYWGCGTIVTSSGGSGS 247  
QY 99 -----W----- 99  
DB 248 EVQLLESGGGLVQPGSLKLSCAASGDFPSRYMNSWVRQAPGKLEWGEINPDSTINY 307  
QY 100 -----EVH-----GYPPYGGCGTIVTSSGG 121  
DB 308 TPSSKDRFTISRDNAKNTLYLQMSKVRSEDTALYYCARLQGMGFYDYGCGTIVTSSGG 367  
QY 122 GSGSGSGSGSGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNA-----VAMVQKPKG 175  
DB 368 GSGSGSGSGSGSELVMTQSPSSLTVTNGERVYMSCKSSQSLNSGNKNTLYTQKPKG 427  
QY 176 QSPRLIYASASRYTVGVPFRFTGSGSGDPFTFTISSVQAEADLAIVYFCQOHRFTPTFGSG 235  
DB 428 QPRLIYMASTRSGVDPDRFTGSGSGDFTLTISSVQAEADLAIVYFCQNDVSYPITFGAG 487  
QY 236 TKLEIK 241  
DB 488 TKLEIK 493

RESULT 9  
US-09-766-543-10  
Sequence 10, Application US/09766543  
Patent No. US20020041865A1  
GENERAL INFORMATION:  
APPLICANT: AuscIn, Richard  
APPLICANT: Kwok, Cheuk S.  
TITLE OF INVENTION: METHODS FOR TREATING TUMORS  
FILE REFERENCE: PP01679.002  
CURRENT APPLICATION NUMBER: US/09/766,543  
CURRENT FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: 60/1177,258  
PRIOR FILING DATE: 2000-01-20  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 10  
LENGTH: 267  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: 520C9  
OTHER INFORMATION: humanized single-chain antibody used in the

OTHER INFORMATION: IL-2-antibody fusions  
US-09-766-543-10  
Query Match 38.9%; Score 913; DB 9; Length 267;  
Best Local Similarity 68.4%; Pred. No. 2.3e-56;  
Matches 167; Conservative 37; Mismatches 36; Indels 4; Gaps 2;

QY 1 QIQVQSGPELKKRGETVTKISCKASGYPTNYGMNVRQAQPGQGLKMWGINTSTGESE 60  
DB 23 EMQLVESGPEVKKKCAVTKVSCASGYPTNYGMNVRQAQPGQGLKMWGINTSTGESE 82  
QY 61 ADPFKGRPFSLTSANTAYLQINNLSKSEDMATYFCARWEVYHGVYPYGGCGTIVTSSG 120  
DB 83 ADPFKERVMTTPTSTAYMDLRSLSDDTAIVYCAR---RFGFA-YWGCGTLVTSSG 138  
QY 121 GSGSGSGSGSGSDIQLTQSHKFLSTSVGDRVSITCKASQDVYNAVAMVQKPKG 180  
DB 139 GSGSGSGSGSGSDIQLTQSPSSLSASVGRVITTCASQDIGNSLTWYQKPKG 198  
QY 181 LIYASASRYTVGVPFRFTGSGSGDPFTFTISSVQAEADLAIVYFCQOHRFTPTFGSGTKLEI 240  
DB 199 LIYATSLDSGVPSRFGSGSGDFTFTISSLPEDATAYYCLOAYLFPYTFGGGTRLEI 258  
QY 241 KALE 244  
DB 259 KGSE 262

RESULT 10  
US-10-127-890-148  
Sequence 148, Application US/10127890  
Publication No. US20030166196A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnick, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrew, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th Floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10/127,890  
FILING DATE: 23-Apr-2002  
CLASSIFICATION: <Unknown>  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/646,360  
FILING DATE: 13-MAY-1996  
APPLICATION NUMBER: PCT/US94/05348  
FILING DATE: 12-MAY-1994  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992  
APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 200-70.P4  
TELECOMMUNICATION INFORMATION:

TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
TELEX: 650 388-1248  
INFORMATION FOR SEQ ID NO: 148:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 148:  
US-10-127-880-148

Query Match 38.5%; Score 903.5; DB 14; Length 240;  
Best Local Similarity 69.3%; Pred. No. 9,4e-56;  
Matches 167; Conservative 32; Mismatches 41; Indels 1; Gaps 1;

QY 1 QIOLVSGPELKKRGETVYKISCKASGYPTNYGMNVKQAPGGGLKMMGMINTSGESTF 60  
DB 1 EIQLVSGGGLVPRGSGVRSICASGYPTNYGMNVKQAPGGGLKMMGMINTHTGEPT 60  
QY 61 ADFFKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEVHYGVYMGQGTIVYSSG 120  
DB 61 ADFFKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEVHYGVYMGQGTIVYSSG 119  
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAYQKPGSPKL 180  
DB 120 GGGSGGGSGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAYQKPGSPKL 179  
QY 181 LIYASASRYTVGVSRRFTSGSGSDPTFTTSSVQAEADLAVYFCQGHRTPTFGSGTKLEI 240  
DB 180 LIYANRLBSGVSRFSGSGSDPTFTTSSVQAEADLAVYFCQGHRTPTFGSGTKLEI 239  
QY 241 K 241  
DB 240 K 240

RESULT 11  
US-10-717-243-148  
Sequence 148, Application US/10717243  
Publication No. US20050054835A1  
GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.  
Studnika, Gary M.  
TITLE OF INVENTION: Immunotoxins Comprising Ribosome-Inactivating  
Proteins  
NUMBER OF SEQUENCES: 169  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.  
STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661

COMPUTER READABLE FORM:  
MEDIUM TYPE: floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/10717,243  
FILING DATE: 18-NOV-2003  
CLASSIFICATION: 530  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US/08/839,765  
FILING DATE: 15-APR-1997  
APPLICATION NUMBER: US 08/425,336  
FILING DATE: 18-APR-1995  
APPLICATION NUMBER: US 08/064,691  
FILING DATE: 12-MAY-1993  
APPLICATION NUMBER: US 07/988,430  
FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901,707  
FILING DATE: 19-JUN-1992  
APPLICATION NUMBER: US 07/787,567  
FILING DATE: 04-NOV-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: McNicholas, Janet M.  
REGISTRATION NUMBER: 32,918  
REFERENCE/DOCKET NUMBER: 11022US09/200-70.P3.C3  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 312/707-8889  
TELEFAX: 312/707-9155  
INFORMATION FOR SEQ ID NO: 148:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 240 amino acids  
TYPE: amino acid  
TOPOLOGY: linear  
MOLECULE TYPE: protein  
SEQUENCE DESCRIPTION: SEQ ID NO: 148:  
US-10-717-243-148

Query Match 38.5%; Score 903.5; DB 17; Length 240;  
Best Local Similarity 69.3%; Pred. No. 9,4e-56;  
Matches 167; Conservative 32; Mismatches 41; Indels 1; Gaps 1;

QY 1 QIOLVSGPELKKRGETVYKISCKASGYPTNYGMNVKQAPGGGLKMMGMINTSGESTF 60  
DB 1 EIQLVSGGGLVPRGSGVRSICASGYPTNYGMNVKQAPGGGLKMMGMINTHTGEPT 60  
QY 61 ADFFKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEVHYGVYMGQGTIVYSSG 120  
DB 61 ADFFKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEVHYGVYMGQGTIVYSSG 119  
QY 121 GGGSGGGSGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAYQKPGSPKL 180  
DB 120 GGGSGGGSGGGSDIQLTQSHKFLSTVGDVRSITCKASQDVYNAVAYQKPGSPKL 179  
QY 181 LIYASASRYTVGVSRRFTSGSGSDPTFTTSSVQAEADLAVYFCQGHRTPTFGSGTKLEI 240  
DB 180 LIYANRLBSGVSRFSGSGSDPTFTTSSVQAEADLAVYFCQGHRTPTFGSGTKLEI 239  
QY 241 K 241  
DB 240 K 240

RESULT 12  
US-09-766-543-12  
Sequence 12, Application US/09766543  
Patent No. US20020041865A1  
GENERAL INFORMATION:  
APPLICANT: Austin, Richard  
APPLICANT: Kwok, Cheuk S.  
APPLICANT: Ring, David B.  
TITLE OF INVENTION: METHODS FOR TREATING TUMORS  
FILE REFERENCE: PP01679,002  
CURRENT APPLICATION NUMBER: US/09/766,543  
PRIOR FILING DATE: 2000-01-19  
PRIOR APPLICATION NUMBER: 60/177,258  
NUMBER OF SEQ ID NOS: 14  
SOFTWARE: PatentIn Ver. 2.0  
SEQ ID NO 12

LENGTH: 276  
TYPE: PRT  
ORGANISM: Artificial Sequence  
FEATURE:  
OTHER INFORMATION: Description of Artificial Sequence: H520C9sfv plus  
US-09-766-543-12

Query Match 38.3%; Score 897; DB 9; Length 276;  
Best Local Similarity 69.7%; Pred. No. 3.1e-55;

Matches 168; Conservative 34; Mismatches 35; Indels 4; Gaps 2;

```

Qy 1 QIOLVOSGPELKKPGETVKISCKASGYPTNYGNMVKQAPQGLKMMGWINTSGESTF 60
Db 23 EIQLVOSGPEVKKPKGASVKISCKASGYTFANYNMVKQAPQGLKMMGWINTYGGSTY 82
Qy 61 ADPFKGRFDSLETSANTAYIQINNLSKSEDMATTFCAWMEVYHGVYVPGGTTVTYSSG 120
Db 83 ADPFKERTFTLDTSTFAHLEISLSREDDATYFCAR--RFGFA-YMGGTLVTYSSG 138
Qy 121 GGGGGGGGGGGGGGGGDIQLOSHKFLSTSVGDRVSTTCASQDVYNAVAYQKPGQSPKL 180
Db 139 GGGGGGGGGGGGGGGGDIQLOSHKFLSTSVGDRVSTTCASQDVYNAVAYQKPGQSPKL 198
Qy 181 LIYASASRYTGVSPRFTGSGSGPDFTFTISSVQAEADLAVYFCQOHFRTPFTFGSGTKLEI 240
Db 199 LIYATSSLDSPSRFSGSGSGDTTLTSSLDQPEDFVYVYCLQYALFPYFGGQTKLEI 258
Qy 241 K 241
Db 259 K 259

```

RESULT 13

```

US-10-071-485-85
; Sequence 85, Application US/10071485
; Publication No. US20030099648A1
; GENERAL INFORMATION:
; APPLICANT: Buyee, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC
; TITLE OF INVENTION: SHOCK,
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015
; CURRENT APPLICATION NUMBER: US/10/071.485
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 09/485,737
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 85
; LENGTH: 541
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: SYNTHETIC
US-10-071-485-85

```

Query Match 37.9%; Score 889.5; DB 14; Length 541;  
 Best Local Similarity 67.7%; Pred. No. 2.3e-54;  
 Matches 168; Conservative 30; Mismatches 47; Indels 3; Gaps 2;

```

Qy 1 QIOLVOSGPELKKPGETVKISCKASGYPTNYGNMVKQAPQGLKMMGWINTSGESTF 60
Db 1 QVQVVOGSGELKKRKGASVKISCKASGYTFDYGMMVKQAPQGLKMMGWINTYGSSTY 60
Qy 61 ADPFKGRFDSLETSANTAYIQINNLSKSEDMATTFCAWMEVYHGVYVPGGTTVTYSSG 120
Db 61 VDDPKGRFVPSLDTSVSAAYLQISLSLKAEADATYFCARGFY--AMDYWGQTTVTYSSG 118
Qy 121 GGGGGGGGGGGGGGGGDIQLOSHKFLSTSVGDRVSTTCASQDVYNAVAYQKPGQSPKL 180
Db 119 GGGGGGGGGGGGGGGGDIQLOSHKFLSTSVGDRVSTTCASQDVYNAVAYQKPGQSPKL 177
Qy 181 LIYASASRYTGVSPRFTGSGSGPDFTFTISSVQAEADLAVYFCQOHFRTPFTFGSGTKLEI 240
Db 178 LIYDTSNLAVGVPARFSGSGSGTSLTISLMEBEDATYFCQSSSYPTFFGGQTKLEI 237

```

```

Qy 241 KALEISNS 248
Db 238 KRTPLGDT 245

```

RESULT 14

```

US-10-071-485-2
; Sequence 2, Application US/10071485
; Publication No. US20030099648A1
; GENERAL INFORMATION:
; APPLICANT: Buyee, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC
; TITLE OF INVENTION: SHOCK,
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015
; CURRENT APPLICATION NUMBER: US/10/071.485
; PRIOR FILING DATE: 2002-02-07
; PRIOR APPLICATION NUMBER: 09/485,737
; PRIOR FILING DATE: 2000-02-14
; PRIOR APPLICATION NUMBER: PCT/EP 98/05165
; PRIOR FILING DATE: 1998-08-14
; PRIOR APPLICATION NUMBER: EPO 98870139.7
; PRIOR FILING DATE: 1998-06-18
; PRIOR APPLICATION NUMBER: EPO 97870122.5
; PRIOR FILING DATE: 1997-08-18
; NUMBER OF SEQ ID NOS: 104
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 2
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Synthetic
US-10-071-485-2

```

Query Match 37.9%; Score 888.5; DB 14; Length 267;  
 Best Local Similarity 69.7%; Pred. No. 1.2e-54;  
 Matches 168; Conservative 27; Mismatches 43; Indels 3; Gaps 2;

```

Qy 1 QIOLVOSGPELKKPGETVKISCKASGYPTNYGNMVKQAPQGLKMMGWINTSGESTF 60
Db 23 QVQVVOGSGELKKRKGASVKISCKASGYTFDYGMMVKQAPQGLKMMGWINTYGSSTY 82
Qy 61 ADPFKGRFDSLETSANTAYIQINNLSKSEDMATTFCAWMEVYHGVYVPGGTTVTYSSG 120
Db 83 VDDPKGRFVPSLDTSVSAAYLQISLSLKAEADATYFCARGFY--AMDYWGQTTVTYSSG 140
Qy 121 GGGGGGGGGGGGGGGGDIQLOSHKFLSTSVGDRVSTTCASQDVYNAVAYQKPGQSPKL 180
Db 141 GGGGGGGGGGGGGGGGDIQLOSHKFLSTSVGDRVSTTCASQDVYNAVAYQKPGQSPKL 199
Qy 181 LIYASASRYTGVSPRFTGSGSGPDFTFTISSVQAEADLAVYFCQOHFRTPFTFGSGTKLEI 240
Db 200 LIYDTSNLAVGVPARFSGSGSGTSLTISLMEBEDATYFCQSSSYPTFFGGQTKLEI 259
Qy 241 K 241
Db 260 K 260

```

RESULT 15

```

US-10-071-485-90
; Sequence 90, Application US/10071485
; Publication No. US20030099648A1
; GENERAL INFORMATION:
; APPLICANT: Buyee, Marie-Ange
; APPLICANT: Sablon, Erwin
; TITLE OF INVENTION: INTERFERON-gamma-BINDING MOLECULES FOR TREATING SEPTIC
; TITLE OF INVENTION: SHOCK,
; TITLE OF INVENTION: CACHEXIA, IMMUNE DISEASES AND SKIN DISORDERS
; FILE REFERENCE: INNS:015

```



GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 20, 2005, 10:01:29 ; Search time 5925.62 Seconds  
(without alignments)  
3622.520 Million cell updates/sec

Title: US-09-596-774-7  
Perfect score: 2345  
Sequence: 1 QIOLVQSGPRLKKRGFTVKI.....LSTATKDTALHMQTLAPR 443

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues  
Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+ p2n.model -DEV=xlh  
-Q=/cgnt2\_1/USPTO.spool\_h/US09596774/runat.20042005.075442.12508/app\_query.fasta\_1.1230  
-DB=GenEmbl -QFMT=faclap -SUFFIX=apr20.rge -MINMATCH=0.1 -LOOPCL=0 -LOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45  
-DOCALLIGN=200 -THR SCORE=spc -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTFMT=prc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USRR=US09596774.QCGN.1.1.6663@runat.20042005.075442.12508 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-Fgapext=7 -Ygapop=10 -Ygapext=0.5 -Delop=6 -Delext=7

Database :

GenEmbl:\*  
1: gb\_da:\*  
2: gb\_htg:\*  
3: gb\_in:\*  
4: gb\_cm:\*  
5: gb\_ov:\*  
6: gb\_pac:\*  
7: gb\_ph:\*  
8: gb\_pl:\*  
9: gb\_pr:\*  
10: gb\_ro:\*  
11: gb\_sts:\*  
12: gb\_sy:\*  
13: gb\_un:\*  
14: gb\_vl:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2345	100.0	1479	6	A47281 Sequence 5
2	1501.5	64.0	1284	6	A93938 Sequence 3
3	1501.5	64.0	3058	6	A93936 Sequence 1
4	1312.5	56.0	2070	6	A57337 Sequence 7

5	1312.5	56.0	2730	6	A57353 Sequence 23
6	1310.5	55.9	2012	6	A22539 Sequence 13
7	1310.5	55.9	2012	6	A22539 Sequence 13
8	1308	55.8	1862	6	A50995 Sequence 36
9	1308	55.8	1862	6	A268367 Sequence 34
10	1308	55.8	1919	6	A50993 Sequence 34
11	1308	55.8	1919	6	AR268366 Sequence 11
12	1308	55.8	2214	6	A57341 Sequence 11
13	1308	55.8	2793	6	A57343 Sequence 21
14	1308	55.8	2793	6	A57351 Sequence 21
15	1308	55.8	3177	6	A57347 Sequence 17
16	1305	55.7	1692	6	A50960 Sequence 1
17	1305	55.7	1692	6	AR268336 Sequence 1
18	1303	55.6	894	12	XXU64991
19	1303	55.6	2233	6	A22470 M.musculus/
20	1303	55.6	2233	6	I28521 Sequence 6
21	1302	55.5	794	6	A57333 Sequence 3
22	1302	55.5	2718	6	A57349 Sequence 19
23	1302	55.5	2793	6	A57345 Sequence 15
24	1285.5	54.8	748	6	A22469 M.musculus
25	1285.5	54.8	748	6	A47277 Sequence 1
26	1285.5	54.8	748	6	I28520 Sequence 4
27	1013	43.2	726	6	AX100174 Sequence 4
28	1011.5	43.1	1422	6	A63768 Sequence 32
29	1011.5	43.1	1545	6	A63770 Sequence 34
30	994	42.4	771	6	AX057978 Sequence 38
31	981.5	41.9	909	6	AR007979 Sequence 1
32	981.5	41.9	909	6	AR058994 Sequence 1
33	981.5	41.9	909	6	I23444 Sequence 1
34	981.5	41.9	909	6	I43364 Sequence 7
35	981.5	41.9	909	6	I61419 Sequence 7
36	981.5	41.9	909	6	I96033 Sequence 7
37	975.5	41.6	777	6	AX935402 Sequence 7
38	975.5	41.6	987	6	AX935453 Sequence 38
39	962.5	41.0	1956	6	A63774 Sequence 40
40	962.5	41.0	2079	6	A63776 Sequence 40
41	960.5	41.0	777	6	AX935404 Sequence 40
42	960.5	41.0	987	6	AX935455 Sequence 40
43	942	40.2	1632	12	AY452127 Synthetic
44	942	40.2	1761	12	AY452134 Synthetic
45	942	40.2	3195	12	AY452128 Synthetic

#### ALIGNMENTS

RESULT 1  
A47281 LOCUS A47281 1479 bp DNA linear PAT 07-MAR-1997  
DEFINITION Sequence 5 from Patent WO9530014.  
ACCESSION A47281  
VERSION A47281.1 GI:2301300  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1479)  
AUTHORS Groner,B and Moritz,D.  
TITLE BIFUNCTIONAL PROTEIN, PREPARATION AND USE  
JOURNAL Patent: WO 9530014-A 5 09-NOV-1995;  
CITBA GEIGY AG (CH)  
COMMENT Other publication ZA 9503440 951102  
Other publication AU 2446995 951129.  
FEATURES  
source  
1..1479  
Location/Qualifiers  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
40..1425  
/codon\_start=1  
/product="LEADER-SCRV(FRPS):LYT-2 HINGE-ZETA"  
/protein\_id="CA03050.1"  
/db\_xref="GI:2301301"  
/translation="MAWVWTLFLMAAKVPOIOLVQSGPRLKKRGFTVKISCKASG"

YFTNYGNMVKQAPGQGLKMMWINTSTGSEFADPKRFDSELSANTAYLOIN  
 NLKSEDAIVYCARMEVHYVYPWGGCTVTVSSGGGGGGGGGGGSILOILOS  
 KFLSTVSDRVSITCKASODVYNAVWQKRGOSPKLITASRRYGVSRRTGSG  
 SQPDFTTSSVOAEDLAVYFCOCHFRPPTFGSSTKLEITALEISNVSMTFSVVP  
 LQKNTSTTKYVLTPTSPVHPTGTSQPORPEDCRPSVKIGDLPDLKCYLDD  
 ILFTIGVITLALYLNKFSRSALNTANIQDNPQLYNEILRREBYDLEKRRADPE  
 MGGKQRRNPQEGVYNALQDKMAEVAEISIGTKERRRGHDLXGLSTATDITY  
 DALHMQITLAPR"

## ORIGIN

## Alignment Scores:

Pred. No.: 4,86e-174 Length: 1479  
 Score: 2345.00 Matches: 443  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 DB: Gaps: 0

US-09-596-774-7 (1-443) x A47281 (1-1479)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuValGlyProGluValThrValIle 20  
 Db CAGATCCAGTGTGTCAGTCTGACCTGAGCTGAAGAAGCCTGAGAGACAGTCAAGATC 153  
 QY 21 SerCysValSerGlyTyrProPheThrAsnTyrGlyMetAsnTTPValIleGlnAla 40  
 Db TCCGCGAAGGCGCTCGGCTATCTTTCAAACTATGAGATGAACTGGGTGAACAGGCT 213  
 QY 41 ProGlyGlnGlyLeuValTyrMetGlyTyrIleAsnThrSerThrGlyIleSerThrPhe 60  
 Db CAGAGACAGGCTTAAAGTGAAGGCGCTGAGTAAACCTCAGCTGAGAGATCAACATTT 273  
 QY 61 AlaAspAspPheValSerGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80  
 Db GCTGATGACTTCAAGGAGCGGTTTGACTTCTTTGAAACCTTCCCAACACTGCTAT 333  
 QY 81 LeuGlnIleAsnLeuValSerGlyLeuAspMetAlaThrTyrPheCysAlaArgTyrGlu 100  
 Db TTGAGAGATCAACCACTCAAAAGTGAAGACATGCTCATATTTCTGTGCAAGTGGAG 393  
 QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 120  
 Db GTTATCAACGCGCTACGTTCTTACTGCGGCGCAAGGACCAAGGACCGGTTCTCTGCG 453  
 QY 121 GlyIleGlySerGlyGlyIleGlySerGlyGlyIleGlySerAspIleGlnLeuThrGln 140  
 Db GGTGGCGGCTTGTGGCGGCGGTGCTCCGCGCGGTGCTGACATCCAGTCAACCGAG 513  
 QY 141 SerHisValPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSer 160  
 Db TCTCACAAAATTCCTGTCCACTTCACTGAGGAGACAGGCTCAGCATCCCAAGGCGAGT 573  
 QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnIleValProGlyIleSerProValLeu 180  
 Db CAGATGTGTATATGCTGTGCTGCTGTATCAACAGAAACAGACAAATCTCTTAACCTT 633  
 QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
 Db CTGATTTACTCGGCATCTCCCGGTACACAGAGATCCCTTCTGCTTCACTGGAGAGGCG 693  
 QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220  
 Db TCTGGGCGGAGTTTCATCTTCAACATCAGAGAGTGTGAGGCTGAAAGCCTGGAGATTTAT 753  
 QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrIleValLeuGluIle 240  
 Db TTCTGTGAGCAACTTTTCTGATCTCATCTCACTGCGCTCGGGGACAAATTTGGAGATC 813  
 QY 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 260  
 Db AAGCTCTAGAGATCAACCACTCGGTATGATCTTCTGTGCGGACAGTCTTCAAG 873  
 QY 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280

Db 874 AAGTGAACCTTACTACTTACCAACCGAGTGTGGAACTCCCTCACTGTGACCTTACC 933  
 QY 281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgIleSerValIleValGlyThr 300  
 Db GGGACATCTAGCCGCCAGAACCAAGAAATTTGTTCGGCCCGGTGCTCAGTGAAGGGAGCC 993  
 QY 301 GlyLeuAspPheLeuGluAspProValLeuCysTyrIleLeuValIleLeuPheIle 320  
 Db GGATTTGACCTTTCTAGAGATCCCAACTCTGCTACTGTAGATGAATCTTCTTATC 1053  
 QY 321 TyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340  
 Db TACGAGTCAATCATCAACAGCCCTGTACTCTGAGAGCAAAATTTCAAGAGTGCAGAGACT 1113  
 QY 341 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlyLeuAsnLeuGlyIleValArgGlu 360  
 Db GCTGCCAACCTGACGAGCCCAACCAAGCTTCAATGAGCTCAATCTAGGGCGAAGAGAG 1173  
 QY 361 GlyTyrAspValLeuGluValLysValAspArgAlaArgAspProGluMetGlyIleValGln 380  
 Db GAATATGACGCTTTGGAGAGAGAGCGGCTCGGATCTCAAGATGGAGGCAACAGCAG 1233  
 QY 381 ArgArgArgAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlu 400  
 Db AGAGAGAGAACCCCAAGAGGCGTATCAATGCACTGAGAGAAAGACAAGATGGCAGAA 1293  
 QY 401 AlaTyrSerGluIleGlyThrIleValGlyValArgArgArgGlyIleValIleValLeu 420  
 Db GCCTACAGTGAAGTCCGCAAAAGGCGAGAGCGGAGAGGCAAGGGGACAGATGGCCTT 1353  
 QY 421 TyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThrLeu 440  
 Db TACCAAGGTCTCGACACTGCGACCAAGACACCTATGATGCTGCTGATGACAGCCTTG 1413  
 QY 441 AlaProArg 443  
 Db 1414 GCCCTCTGCG 1422

## RESULT 2

A93938 1284 bp DNA linear PAT 22-JUN-2000  
 LOCUS A93938  
 DEFINITION Sequence 3 from Patent WO9720938.  
 ACCESSION A93938  
 VERSION A93938.1 GI:6742040  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 unclassified.

REFERENCE  
 1 (bases 1 to 1284)  
 Rosenthal, P. and Kulmburg, P.  
 CELLS WITH HYBRID RECEPTOR AND GENE CONSTRUCT WHICH CAN BE  
 CONTROLLED BY SAID HYBRID RECEPTOR, AND USE OF SAID CELLS IN GENE  
 THERAPY  
 Patent: WO 9720938-A 3 12-JUN-1997,  
 ROSENTHAL FELICIA (DE); KULMBURG PETER (DE)  
 location/Qualifiers  
 1..1284  
 /organism="unidentified"  
 /mol\_type="unassigned DNA"  
 /db\_xref="taxon:32644"

## FEATURES

source

## ORIGIN

## Alignment Scores:

Pred. No.: 4,08e-108 Length: 1284  
 Score: 1501.50 Matches: 297  
 Percent Similarity: 76.29% Conservative: 44  
 Best Local Similarity: 66.44% Mismatches: 79  
 Query Match: 64.03% Indels: 27  
 DB: Gaps: 6

US-09-596-774-7 (1-443) x A93938 (1-1284)



```

Oy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysProGlyGluThrValLysIle 20
Db 13 CAGGTCCAACTGACAGCTCTGGGGCTGATGTAAGAGCTGGGGCTTCACTGAAGCTG 72
Oy 21 SerGlyValAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db 73 TCCTGCAGAGGCTTCTGGCTACACCTTCAACACCTAGTATGCTGGTGTGAAGCAGAG 132
Oy 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
Db 133 CTGGACAGAGGCTTGAAGTGAATGTAAGATTGATCTCAATATAGTGGTGTACTAAGTAC 192
Oy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 193 AATGAGAAGTTCAGAGAGAGGCCACACTGACTAGACAAACCTCCAGCAGCAGCTTAC 252
Oy 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db 253 ATGCAGCTCAGACCTGACATCTGAGACTCTGGGGTCTATTATTTGTGCAAGATACGAT 312
Oy 101 ValTyr--HisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSer 119
Db 313 TACTACGAGTAGTACTGACTTGTGAGGACCAAGGACCAAGCTACCGCTCTCTCTCA 372
Oy 120 GlyGlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThr 139
Db 373 GGTGAGGCGGTTCAGGGGAGGTGGCTCTGGCGGTGGAGTCTCAGGCTGTGGACA 432
Oy 140 GlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAla 159
Db 433 CAG--GAATCTGCACTCACCACTCACCCTGTGAAACAGTCACACTGCTGTGCTCA 489
Oy 160 SerGlnAspValTyr-----AsnAlaValAlaTrpTyrGlnGlnLysProGlyGln 176
Db 490 AGTACTGGGGCTTTACAACACTAGTACTAGTCAACCTGGGCTCAAGAAACCAATCAT 549
Oy 177 SerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPhe 196
Db 550 TTATTCACTGGCTCTAATAGTGTGATCAACAACAGAGCTCCAGGCTTCTCTCCGATTC 609
Oy 197 ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 216
Db 610 TCAGGCTCCCTGATTTGAGAGACAGAGCTCCCTCACTACACAGGAGGAGCTGAGAT 669
Oy 217 LeuAlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThr 236
Db 670 GAGGCAATATATTTCTGTCTCTATGTAGCAAGCAACATTTGGGTGTGGTGAGAGAAC 729
Oy 237 LysLeuGlnIleLysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValVal 256
Db 730 AAATCTGACTCTC-----CTCGAG----- 747
Oy 257 ProValLeuGlnLysValAlaAsnSerThrThrLysProValLeuArgThrProSerPro 276
Db 748 -----GAATTCACTACTACCAAGCCAGCTGCTCGAATCCCTCACCT 789
Oy 277 ValHisProThrGlyThrSerGlnProGlnArgProGluAspArgProArgGlySer 296
Db 790 GTGACCTCTACCGGAGCATCTCAAGCCCAAGAGATGTCGGCCCGGTGGCTCA 849
Oy 297 ValLysGlyThrGlyLysAspPheLeuGlnLysAspProLysLeuCysTyrLeuLeuAspGly 316
Db 850 GTGAAGGGGACCGATTTGACTTC-----GACCCCAACTCTGTACTTCTGATAGTGA 903
Oy 317 IleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArg 336
Db 904 ATCTCTTCATCTACGAGATCATCATCAAGCCCTGACTGAGAGCAAAATTTCAGCAGG 963
Oy 337 SerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeu 356
Db 964 AGTCAGAGAGCTGCTGCCAACCTGAGAGACCCCAACGAGCTTCAAAATGAGCTCAATCTA 1023
Oy 357 GlyArgArgGlnGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnLysGly 376

```

```

Db 1024 GGGCGAAGAGAGATATGAGCTTTGGAGAAGACCGGGCTCGGGATCCAGAGATGGGA 1083
Oy 377 GlyLysGlnGlnArgArgArgAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAsp 396
Db 1084 GGCAGAACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1143
Oy 397 LysMetAlaGlnAlaTyrSerGlyLysGlyThrLysGlyGlnArgArgGlyLysGly 416
Db 1144 AACATGCGAAGAGCTTACAGTGAATCGGCACAAAGCCGAGAGCGAGAGAGAGAGAG 1203
Oy 417 HisAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHis 436
Db 1204 CACGATGCGCTTTACAGAGGTCTCAGACTCCACCAAGAGACACCTATGATGCCCTCAT 1263
Oy 437 MetGlnThrLeuAlaProArg 443
Db 1264 ATGCAGACCTTGGCCCCCTGCG 1284

RESULT 3
LOCUS A93936 3058 bp DNA linear PAT 22-JAN-2000
DEFINITION Sequence 1 from Patent WO9720938.
ACCESSION A93936
VERSION A93936.1 GI:6742038
KEYWORDS
SOURCE
ORGANISM
unidentified
unclassified.
REFERENCE
1 (bases 1 to 3058)
AUTHORS
Rosenthal, F. and Kulmburg, P.
TITLE
CELLS WITH HYBRID RECEPTOR AND GENE CONSTRUCT WHICH CAN BE
CONTROLLED BY SAID HYBRID RECEPTOR, AND USE OF SAID CELLS IN GENE
THERAPY
JOURNAL
Patent: WO 9720938-A 1 12-JUN-1997;
ROSENTHAL, FELICIA (DE); KULMBURG, PETER (DE)
FEATURES
source
1..3058
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"

ALIGNMENT Scores:
Pred. No.: 1,21e-107 Length: 3058
Score: 1501.50 Matches: 297
Percent Similarity: 76.29% Conservative: 44
Best Local Similarity: 66.44% Mismatches: 79
Query Match: 64.03% Indels: 27
Gaps: 6

US-09-596-774-7 (1-443) x A93936 (1-3058):
Oy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysProGlyGluThrValLysIle 20
Db 862 CAGGTCCAACTGACAGCTCTGGGGCTGATGTAAGAGCTGGGGCTTCACTGAAGCTG 921
Oy 21 SerGlyValAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db 922 TCCTGCAGAGGCTTCTGGCTACACCTTCAACACCTAGTATGCTGGTGTGAAGCAGAG 981
Oy 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
Db 982 CTGGACAGAGGCTTGAAGTGAATGTAAGATTGATCTCAATATAGTGGTGTACTAAGTAC 1041
Oy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnLysAspThrAlaAsnThrAlaTyr 80
Db 1042 AATGAGAAGTTCAGAGAGAGGCCACACTGACTAGACAAACCTCCAGCAGCAGCTTAC 1101
Oy 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db 1102 ATGCAGCTCAGACCTGACATCTGAGAGACTCTCGGTCTATTATTTGTGCAAGATACGAT 1161

```

```

QY 101 ValTYR---HisGLYTYRValProTYRTPGlyGlnGlyThrThrValThrValSerSer 119
DB 1162 TACTACGGAGTAGTACTTCTTACCTACCTGGGCCAAGGACCAAGCTACCTCTCTCTCA 1221
QY 120 GtlyGlyGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 139
DB 1222 GGTGAGGCGCGGTTCAGGCGAGGTGCTCTGGCGGTGCGGATCTCAGGCTTTGGGACA 1281
QY 140 GlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerHisLysCysLysAla 159
DB 1282 CAG---GAATCTGCACATCCACATACCTGCTGTAAGAGTCACTCACTTCTGCTCA 1338
QY 160 SerGlnAspValTYR-----AsnAlaValAlaTyrTYRglnGlnLysProGlyGln 176
DB 1339 AGTACTGGGGCTGTACAACTAGTAACTATGCACTGGGTCCAGAAAGAAAACCATATCAT 1398
QY 177 SerProLysLeuLeuLysLysSerHisSerValTYRThrGlyValProSerArgPhe 196
DB 1399 TTATTCACCTGGCTTAATGAGTGTGTACCAACCAACGAGCTCCAGGTCTCTGCGCATTC 1458
QY 197 ThrGlySerGlySerGlyProAspPheThrPheThrLysSerSerValGlnAlaGlnAsp 216
DB 1459 TCAGGCTCTCCGATTTGGAGACAGGCTGCTCCACATCAAGGCGGACAGACTGGAGAT 1518
QY 217 LeuAlaValTYRpheCysglnGlnHisPheArgThrProPheThrPheGlySerGlyThr 236
DB 1519 GAGGCAATATATTTCTGTCTCTATGTACAGCAACCATTTGGGTGCTGGTGGAGGAACC 1578
QY 237 LysLeuGlnLysLysLysLeuGlnLysSerAsnSerValMetTYRpheSerSerValVal 256
DB 1579 AAACGACTGTC-----CTCGAG----- 1596
QY 257 ProValLeuGlnLysValAsnSerThrThrLysProValLeuArgThrProSerPro 276
DB 1597 -----GAATTCACCTAATCTCAACAGGAGCTGCGCAACTCTCTCACT 1638
QY 277 ValHisProThrGlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySer 296
DB 1639 GTGCACCTCACTCGGAGACATCTCAGCCCCAGAGACAGAAAGATTGCGGCCCGGCTCA 1698
QY 297 ValLysGlyThrGlyLysLeuAspPheLeuGlnAspProLysLeuCysTYRLeuLeuAspGly 316
DB 1699 GTGAGGAGGACCGGATGTGACTTC-----GACCCCAAACTCTCTCTGCTGATGCA 1752
QY 317 IleLeuPheIleTYRglnValIleIleThrAlaLeuTYRLeuArgAlaLysPheSerArg 336
DB 1753 ATCTCTTCTCATCTACGAGTATCATCATCAGCCCTGTACTGAGACGAAATTCAGACAG 1812
QY 337 SerAlaGlnThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTYRAsnGlnLeuAsnLeu 356
DB 1813 AGTCACAGAGACTGCTGCCAACCTGCAGGACCCCAACGAGCTCTCAATGAGCTCAATCTA 1872
QY 357 GlyArgArgGlnGlyTYRAspValLeuGlnLysLysArgAlaAspProGlnLysMetGly 376
DB 1873 GGGGGAAGAGAGAAATGACGTCTTGGAGAAAGGCGGCTCGGATCCAGAAATGGGA 1932
QY 377 GtlyLysGlnGlnArgArgAspProGlnGlnGlyValTYRAsnAlaLeuGlnLysAsp 396
DB 1933 GCGAAACGACGAGAGAGAGAACCCCGAGAGAGCATATACATGACCTGCGAAGAAAGAC 1992
QY 397 LysMetAlaGlnAlaTYRSerGlnIleGlyThrLysGlyLysArgArgGlyLysGly 416
DB 1993 AACATGCGACAGAGCCTACAGTGAATCGGCACAAAGGCGAGGCGAGGCGAAGGG 2052
QY 417 HisAspGlyLeuTYRglnGlyLeuSerThrAlaThrLysAspThrTYRAspAlaLeuHis 436
DB 2053 CACATGAGCCTTACAGAGGCTTGAAGACTGCGCACCAAGAGACCTATGATGCTCGCAT 2112
QY 437 MetGlnThrLeuAlaProArg 443
DB 2113 ATGCAGACCTCGGCGCTCTCGC 2133

```

RESULT 4

```

A57337
LOCUS A57337 2070 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 7 from Patent EP0739984.
ACCESSION A57337
VERSION A57337.1 GI:3713216
KEYWORDS
SOURCE
ORGANISM
synthetic construct
other sequences/ artificial sequences.
REFERENCE
1. Wels, W.D., Schmidt, M. and Groner, B.D.
Bivalent polypeptides containing at least two domains
Patent: EP 0739984-A 7 30-OCT-1996;
SAN TUMORFORSCHUNGS GMBH (DE)
Location/Qualifiers
1. 2070
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="pSW202-5"
1. 1956
/note="unnamed protein product"
/codon_start=1
/transl_table=11
/protein_id="CAA03441.1"
/db_xref="GI:3713217"
/translation="MKKTAIAIYALVLAAGPATYVAQADYKDDDKLHHHHHKKLOVOLOQ
SGPELKKPGETVAKISKASGYFTNTGMMVAVKQAPQGLKMMGMLNTSGESTFADF
KGRFDFSLSTANTAYLQIINLKSEDMATYFCARMVHYGVYWGQGTIVVSQGG
GGGGGGGGSDIQLTQSHKELSTGVSITCKASODVYNAVVMYQKQSGFKL
LIYSASRYATGVPSRFTGSGSGDPFTTSSVOAEDLAVYPCOHPRTFTGSGFKL
EIKALSHHHHHLEGGSLAALTAHQAACHLPLEFTTHRQPRGMEOLECGGYVORIVA
LYVAELSNQNDQVTRNMLASPGSGDGEAIRBPBQARLALTLIAAESRPARG
TGNDERGANADVSLITCYVAAGECAGPDSGDLUBERYPTGAELFGSDVSESTR
GTQNTVERLLQHRQLBERGYVFGVYHOTFLEAQSIVFGVRAISODLDIRFV
IADPALAAGYADQDEPDARGRINQALIRVYVPSLPGFYRTSLTLAAPAAAGVE
RLIGHPLPLRLDITGPEEGEGLRTILGLMPLAERTVVIIPSAIPTDPRNVGDLDPSS
IPKQEAISALPDYASQPKPFREDLK"
ORIGIN
Alignment Scores:
Pred. No.: 4,56e-93 Length: 2070
Score: 1312.50 Matches: 284
Percent Similarity: 68.59% Conservative: 13
Best Local Similarity: 65.59% Mismatches: 55
Query Match: 55.97% Indels: 83
DB: 6 Gaps: 8
US-09-596-774-7 (1-443) x A57337 (1-2070)
QY 1 GlnIleGlnLeuValGlnSerGlyProGlyLeuLysLysProGlyGlyThrValLysIle 20
DB 115 CAGGTACAACTGCAGCAGCTGTGAGCTGAACCTGAAAGAGCTGGAGAGACAGTCAAGATC 174
QY 21 SerCysLysAlaSerGlyTYRProPheThrAsnTYRglnLysMetAsnTrpValLysGlnAla 40
DB 175 TCTGCAAGGCGCTCTGGATTCCTTCAACAACATAGGAATGAACCTGGGTGAAGCAGCGT 234
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTYRPIleAsnThrSerThrGlyGlnSerThrPhe 60
DB 235 CCAGGACAGAGGTTTAAAGTGATGGGCTGATTAACCTCCACCTGAGAGAGTCAACATT 294
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTYR 80
DB 295 GCTGATGACTTCAAGGAGCGGTTGACTTCTCTTGGAAACCTTGCCAAACCTGCTAT 354
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyLysPheMetAlaThrTYRpheCysAlaArgTrpGln 100
DB 355 TTGCAGATCAACAACCTCAAAAGTGAAGCATGGCTACATATTCTGTGCAAGATGGGAG 414
QY 101 ValTYRHisGlyTYRValProTYRTrpGlyGlnGlyThrThrValThrValSerGly 120
DB 415 GTTACCAAGGCTTACTGAGGCGCAAGGAGACCAAGGATCAACGCTTCTCTGCGC 474

```

```

QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
DB 475 GGTGGCGGTTCTGGTGGGGGTGGCTCCGGCGGTGGCTTGCACATCCAGCTGACCAG 534
QY 141 SerHisGlyPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer 160
DB 535 TCTCAAAATTCCTCTCCACTTCAGTAGAGACAGGGTACGATCACCCTGCAAGGCCAGT 594
QY 161 GlnAspValTyrAsnAlaValAlaTyrPyrGlnGlnLeuProGlyGlnSerProGlyLeu 180
DB 595 CAGGATGATGAATGAATGCTGTCTCTGGTATCAACAGAAACCGACAACTCTTAACTT 654
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 655 CTGATTACTCGGCATCTCTCCCGGTACACTGAGTCCCTTCCGCTTCACTGGCAGTGGC 714
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 715 TCTGGGCGGATTTCACTTTCACATTCAGACAGTGTGACAGGCTGAAAGACTGGCAGTTAT 774
QY 221 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrIleLeuGlnIle 240
DB 775 TTCTGTCAAGCAACATTTCTGTACTCAATTCAGATTCGGCTCGGAGCAAAATTTGAGATC 834
QY 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
DB 835 AAAGCTTAGACCA----- 848
QY 261 LysValAsnSerThrThrThrLysAspValLeuArgThrProSerProValHisProThr 280
DB 849 -----CCATCATCAACATCACCTAGA 869
QY 281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValIleGlyThr 300
DB 870 GGG---CGGACAGCTCGGCGGCTGAC-----CGGACACAGCGCTGCCACCTGCCCT 920
QY 301 GlyLeuAspPheLeuGluAspProLysLeuCystrLeuLeuAspGlyIleLeuPheIle 320
DB 921 GGAGACTTTCACCCCTCATCGCCACCGCGGCTGGGACAACTGAGACAGTGGCGCTA 980
QY 321 TyrGlyValIle-IleThrAlaLeuTyrIleuArgAlaLysPheSer----- 335
DB 981 TCCGGTGACGGCGTGTGCTGCTTACCTGCGCGGACGACTGATGAACAGTGCAG 1040
QY 336 -----ArgSerAlaGluThrAla-----AlaAs 343
DB 1041 CCAAGTGATCCGCAACGCCCTGGCCAGCCCGGACGCGCGCACTGGGCGAAGCGAT 1100
QY 343 nLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGluGluTyrAs 363
DB 1101 CCGGACGACCGGACGAGCCCGCTGACCTTGCACCT-GGCCCGCGCGAGAC----- 1154
QY 363 pValLeuGlnLysArgAlaArgAspProGluMetGlyGlyGlnGlnArgArg 383
DB 1155 -----CGAGGCTTCCGTCG-----GCAGGCGACCGGCAACGACGA-- 1190
QY 383 GAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSe 403
DB 1190 ----- 1190
QY 403 rGluIleGlyThrLysGlyLysArgArgGlyLys 415
DB 1191 -----GGCCGGCGGCGCAACCGGACGTCGTAG 1220

```

```

RESULT 5
A57353 2730 bp DNA linear PAT 03-MAR-1998
LOCUS DEFINITION Sequence 23 from Parent EP073984.
ACCESSION A57353 GI:3713232
VERSION A57353.1
KEYWORDS synthetic construct
SOURCE

```

## ORGANISM synthetic construct

other sequences; artificial sequences.

## REFERENCE

## AUTHORS

Wels, W.D., Schmidt, M. and Groner, B.D.

## TITLE

Bivalent polypeptides containing at least two domains

## JOURNAL

Patent: EP 073984-A 23 30-Oct-1996;

## FEATURES

SAN TUDORPORSCHEUNGS GMBH (DE)

## source

Location/Qualifiers

## CDS

```

1..2730
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/clone="PMS246-5-5 KDEL"
1..2700
/notes="unnamed protein product"
/codon_start=1
/transl_table=1
/protein_id="CA03449.1"
/db_xref="GI:371323"
/translation="MKKTAIAIIVALAGFATVAADYKDDDKLHHHHHKLQVQLQ
SGPELKKRGERTVTKSCKASGYPTNYGNNVWKOAPGGLKMGWINTSTGTFADDF
KGRDPSLETSANTRAIQIINLKSEDMATYRCAMEYHGVYPTWGCTTYTSRGGG
GSGGGSGGSGSDIQLTQSHKFLSTVGDVSTICKASQDVNVAWMTQKPGSPKL
LIYASASRYTGPSPFTGSDPFTFTISSVDAEDLAVYPCQHFRTPTFGSCTKL
EIKALEHHHHHLEGGSLAALTAAOACHLPLETFTPRQPRGMEQLQECGYVORLVA
LYLAARLSMNVQDVITRNALASPGSGDLGAIIEOPQARLALTLLAAESERFVROG
TGNDAGAAANDVSLTCCPVACGACGADSGDALLERNYPTGAEFTGDDGVSPSTR
GTQWNTYERLLQARQLRERGYVFGHTGTELEAOSIVFGGVARASQDDIATRGFY
IAGDPALAYGAQDQEPDARGIRNGALIRVYRPSLPGFRTSLTAAEALAEVE
RLIGHPLRLIDATGPEBEGRLLETIGWLAERTVVIPIAIPDPNNVGGDLDPSS
IPDKEOAIISALPDVASQPKPSPASQVQLQSPGLKPEETVAKISKASGYPTNY
GNNVWKOAPGGLKMGWINTSTGTFADDFKRPDPSLETSANTRAIQIINLKSE
MATYRCAMEYHGVYPTWGCTTYTSVSGGGSGGSGSDIQLTQSHKFLSTV
VGDRVSTICKASQDVNVAWMTQKPGSPKLIYASASRYTGPSPFTGSGSDPFT
FTISSVDAEDLAVYPCQHFRTPTFGSCTKLIEKALEKDEL"

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 6,44e-93 Length: 2730
Score: 1312.50 Matches: 284
Percent Similarity: 68.59% Conservative: 13
Best Local Similarity: 65.59% Mismatches: 55
Query Match: 55.97% Indels: 83
DB: 6 Gaps: 8

```

US-09-596-774-7 (1-443) x A57353 (1-2730)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValIle 20
DB 115 CAGGTAAACACTGACGACGAGTGTGACCTGAAGAAAGCTGGAGAGACAGTCAAGTTC 174
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 40
DB 175 TCTCGCAAGGCTCTGGGTATCTTTCACAAACATGAAAGAACTGGGTGAAGAGCGCT 234
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
DB 235 CCAGGACAGGCTTTAAAGTGAAGGCTGATTAACCTTCACTGAGAGTCAACATTT 294
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 295 GCTATGACTTCAAGGACGAGCGTTTGACTTCTTGGAAACCTTGCACAAACAGTCCAT 354
QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
DB 355 TTGCAGATCAACCACTCAAAAGTGAAGACATGCTCATATTTCTGTGCAAGATGGAG 414
QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 120
DB 415 GTTAAACAGGCTACGTTCTTACTGGGCGCAAGGACACAGGTCACCGTTCTCTGGC 474
QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140

```

```

Db 475 GGTGGCGGTTCTGTGGTGGTGGCTCCGGCGGTGGCTTGACATCCAGTCAGCCAG 534
Qy 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCylAspAlaSer 160
Db 535 TCTCAAAATTCCTCTCCATCTTCAGTAGAGACAGGGTCAGATCACTGCAGAGCCAGT 594
Qy 161 GlnAspValTyrAspAlaValAlaIleTyrGlnGlnLysProGlnGlnSerProLysLeu 180
Db 595 CAGATGTGTATTAATGCTGTGGTCCGTATCAACAGAAACAGGCAATCTCTTAATCT 654
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 655 CTGATTACTCGGCATCTCTCCGCTACAGTGGAGTCCCTTCTCGCTTCACTGGCAGTGGC 714
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 220
Db 715 TCTGGGCGGATTTTCACTTTCACTTCACATCAGAGTGGAGGCTGAGAACCTGACGTTAT 774
Qy 221 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
Db 775 TTCTGTCAACAACATTTTCTGTAATTCATTCAGTTCCGCTCGGGGACAAATTTGGAGATC 834
Qy 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 260
Db 835 AAAGCTCTAGAGCA----- 848
Qy 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
Db 849 -----CCATCATCACTCACTCACTCA 869
Qy 281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgLysSerValLysGlyThr 300
Db 870 GGG---CGGCAGCTGGCGCGCTGAC-----CGCGCACAGGCTGGCCACCTGCCGCT 920
Qy 301 GlyLeuAspPheLeuGlnAspProLysLeuLysCysTyrIleLeuAspGlyIleLeuPheIle 320
Db 921 GGAACATTTTCAACCCGCTATCGCCAGCGCGCTGGAGAACACTGGAGACAGTCCGCTA 980
Qy 321 TyrGlyValIle-IleThrAlaLeuTyrLysArgAlaLysPheSer----- 335
Db 981 TCCGCTGACAGGCGTGGTCCCTCTTACCTGGCGCGGACTGATCAAGAACAGGTGCA 1040
Qy 336 -----ArgSerAlaGlnThrAla-----AlaAs 343
Db 1041 CCAAGTATCCGCAACGCCCTCGCCAGCCGCGCAGCGCGCAGCACTGGCGCAAGCAT 1100
Qy 343 nLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlnGlyTyrAs 363
Db 1101 CCGCGACAGCGCGAGCAGGCCCTGTGGCCCTGACCT--GGCCGCGCGCAGAG----- 1154
Qy 363 pValLeuGlnLysLysArgAlaArgAspProGlnMetGlyGlyLysGlnGlnArgArg 383
Db 1155 -----CGAGGCTTGTCTCG-----GCAGGAGCAGCGGACACAGCA-- 1190
Qy 383 GAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSe 403
Db 1190 ----- 1190
Qy 403 rGlnIleGlyThrLysGlyGlnArgArgGlyLys 415
Db 1191 -----GGCCGCGCGCGCAACGCCCACTGTGTAG 1220

```

```

RESULT 6
A22539          2012 bp   DNA       linear   PAT 15-NOV-1994
LOCUS          M.musculus/P.aeruginosa Fv(FRP5)-ETa fusion protein.
DEFINITION
ACCESSION      A22539
VERSION        A22539.1 GI:641549
KEYWORDS
SOURCE         synthetic construct
SYNTHETIC     other sequences; artificial sequences.
ORGANISM
REFERENCE      1 (bases 1 to 2012)

```

```

AUTHORS      Weis, W.S., Hynes, N.E., Harwerth, I.M., Groner, B., Hardman, N. and
              Zwick, M.
TITLE        Recombinant antibodies specific for a growth factor receptor
JOURNAL      Patent: EP 0502812-A 10 09-SEP-1992;
              CIBA-GEIGY AG
FEATURES
source       Location/Qualifiers
              1..2012
              /organism="synthetic construct"
              /mol_type="unassigned DNA"
              /db_xref="taxon:32630"
              1..1911
              /product="Fv heavy and light chain variable domain and
              exotoxin A fusion protein"
ORIGIN
Alignment Scores:
Pred. No.:      6,3e-93      Length:      2012
Score:          1310.50      Matches:     279
Percent Similarity: 68.51%   Conservative: 19
Best local Similarity: 64.14% Mismatches:    43
Query Match:    55.88%      Indels:      95
Db:              6          Gaps:        10
US-09-596-774-7 (1-443) x A22539 (1-2012)
Qy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysLysProGlnGlnThrValLysIle 20
Db 97 CAGGTACACTGACGACAGTGTGACCTGAACCTGAAGAGCTGGAGACAGTCAAGATC 156
Qy 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
Db 157 TCTCGAAGGCTCTGTGGTATCTTTCACAACTATGATGATGAACCTGGGTGAAGCGCT 216
Qy 41 ProGlyGlnGlyLeuLysTyrPheGlyTyrIleAsnThrSerThrGlyLysThrPhe 60
Db 217 CCAGACAGGGTTTAAAGTGAAGGAGGCTGATTAAACCTCCACCTGAGAGTCAACATT 276
Qy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db 277 GCTGATGACTTCAAGGAGCGTTTGACTTCTCTTGGAAACCTGTGCCAACACTGCTTAT 336
Qy 81 LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTyrPheCysAlaArgTyrGln 100
Db 337 TTSCAATCAACAACCTCAAAAGTGAAGCATGGCTACATATTTCTGTCCAAAGATGGAG 396
Qy 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrValThrValSerSerGly 120
Db 397 GTTTACACAGGCTACGTTCTTACTGGGGCAAGGACACAGCTACCGTTTCTCTGCGC 456
Qy 121 GlyGlyLysSerGlyGlyGlyLysSerGlyGlyGlyLysSerAspIleGlnLeuThrGln 140
Db 457 GGTGGCGGTTCTGTGGCGGCTGCGCGGCTGCGGCTGCTTGCATCTCAGCTGACCCAG 516
Qy 457 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCylAspAlaSer 160
Db 517 TCTCAAAATTCCTGTCACCTTCAGTAGAGACAGGTCAGATCACTGCAGAGCCAGT 576
Qy 161 GlnAspValTyrAspAlaValAlaIleTyrGlnGlnLysProGlnGlnSerProLysLeu 180
Db 577 CAGGATGTATTAATGCTGTGGTCCGTATCAACAGAAACAGGCAACATCTCTAACTT 636
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 637 CTGATTACTCGGCATCTCTCCGCTACATGAGTCCCTTCTCGCTTCACTGGCAGTGGC 696
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 220
Db 697 TCTGGGCGGATTTTCACTTCACTACATGAGTGTGCAAGGCTGAAGACCTGGCAGTTAT 756
Qy 221 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
Db 757 TTCTGTCAACAACATTTTCTGTAATTCATTCACGTTCCGCTCGGGGACAAATTTGGAGATC 816

```

Qy 241 LysalaleuGlu-----IleSerAsnSerValMetTyrPheSerSerValPro 257  
Db 817 AAAGCTTAAAGAGGGGCGGAGCGCTGGCCGGCGGACCGGACCAAGCCCTGCACCTGGCG 876  
Qy 258 ValLeuGlnLysValAsnSerThrThrLys----- 268  
Db 877 CTG-----GAGACTTTCACCCCTCATCCGCGACCGCGCGCTGGAGAACAA 921  
Qy 269 -----ProValLeuArgThrProSerProValHisProThrLysThr 282  
Db 922 CTGGAGCAGTCGGCGCTATCCGGTCGACGGCGTGGT-CGCCCTCTA---CTGGCGCGCGCG 977  
Qy 283 Ser-----GlnProGlnArgProGlnuAapCysArgProArgLysValLysGlyThr 300  
Db 978 ACTGTCATGAAACCAAGTCGACAGTGTATCCGCAACCGCCGCGACCGCCGCGAGCGG 1037  
Qy 301 GlysLeuAapPheLeuGlnuAapProLysLeuCystrLeuLeuAapGlyLleLeuPheLle 320  
Db 1038 CGGCGACCTGGGCGGAGCGATCCG----- 1061  
Qy 321 TyrGlyValLleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340  
Db 1062 -----CGAGCAGCCGCGAGCA 1076  
Qy 341 AlAlaAsnLeuGlnuAapProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlu 360  
Db 1077 GGGCCGCTGGCGCCCTGACCT-----GGCGCGCGCGCGA 1109  
Qy 361 GlnTyrAapValLeuGlnLysLysArgAlaArgAapProGlnuAapGlyLysGlnGln 380  
Db 1110 GAG-----CGACCGCTTCGTCGCG-----GCAGGCGACCGCGCA 1142  
Qy 381 ArgArgArgAapProGlnGlnGlyValTyrAsnAlaLeuGlnLysAapLysMetAlaGlu 400  
Db 1143 CGAGCA----- 1148  
Qy 401 AlaTyrSerGlnLleGlyThrLysGlyGlnuArgArgGlyLys 415  
Db 1149 -----GGCCGGCGCGCGCAACCGCCGATGGTGGAG 1178

RESULT 7  
LOCUS 128526 2012 bp DNA linear PAT 06-FEB-1997  
DEFINITION Sequence 13 from patent US 5571894.  
ACCESSION 128526  
VERSION 128526.1 GI:1819302  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 2012)  
AUTHORS Weis,W.S., Hyner,N.E., Harweth,I.-M., Groner,B., Hardman,N. and Zwacki,M.  
TITLE Recombinant antibodies specific for a growth factor receptor  
JOURNAL Patent: US 5571894-A 13 05-NOV-1996;  
FEATURES  
source location/Qualifiers  
1..2012  
/organism="unknown"  
/mol\_type="unassigned DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 6.3e-93 Length: 2012  
Score: 1310.50 Matches: 279  
Percent Similarity: 68.51% Conservative: 19  
Best Local Similarity: 64.14% Mismatches: 43  
Query Match: 55.88% Indels: 95  
DB: Gaps: 10  
US-09-596-774-7 (1-443) x 128526 (1-2012)

Qy 1 GlnLleGlnLeuValGlnSerGlyProGlnuLeuLysLysProGlyGlnThrValLysLle 20  
|||||  
|||||

Db 97 CAGGTAACTGACAGCTCTGACCTGAAGCTGAAGAGCTGGAGACAGTCAAGATC 156  
Qy 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40  
Db 157 TCTGCAAGAGCCCTCGGGTATCTTTCACAAACTATGAAATGAACTGGGTGAAGAGGCT 216  
Qy 41 ProGlyGlnGlyLeuLysTyrPheGlyTyrPheAsnThrSerThrGlyLysThrPhe 60  
Db 217 CCAGAGCAGGGTTTAAAGTGAATGGGCTGGATTAACCTCCACCTGAGAGATCAATTT 276  
Qy 61 AlaAapAapPheLysGlyArgPheAapPheSerLeuGlnuThrSerAlaAsnThrAlaTyr 80  
Db 277 GCTATGACTTCAAGAGGAGCGGTTTGACTTCTTGGAAACCTGCGCAACCTGCGCAT 336  
Qy 81 LeuGlnLleAsnAsnLeuLysSerGlnuAapPheAlaThrTyrPheCysAlaArgTyrGlu 100  
Db 337 TTGCAGATCAACAACCTTAAGATGAACATGGCTATATTTCTGTGCAGATGGAG 396  
Qy 101 ValTyrHisGlyTyrValProTyrTyrPheGlyGlnGlyThrThrValThrValSerSerGly 120  
Db 397 GTTTAACACGGGCTACGTTCTTACCTGGGCGCAAGGACCAAGTCACCGTTCTCTGGC 456  
Qy 121 GlyLysGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerAapLleGlnLeuThrGln 140  
Db 457 GGTGGCGGTTCTGGTGGCGGTGGCTCCGCGCGGTGGCTTGCATCCAGCTGACCGCAG 516  
Qy 141 SerHisLysPheLeuSerThrSerValGlyAapArgValSerIleThrCysLysAlaSer 160  
Db 517 TCTCAACAATTCCTGTCACCTTCACTAGTAGAGACAGGGTCAACATCCTGCAAGGCCCT 576  
Qy 161 GlnAapValTyrAsnAlaValAlaTyrGlnGlnLysProGlyGlnSerProLysLeu 180  
Db 577 CAGATGTGTATTAATGCTGTTGCTGGATCAACAGAAACAGGACAAATCTCTTAATCTT 636  
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
Db 637 CTGATTTAACTCGGATCCTCCCGGTACCTGAGATCCCTTCTCTTCACTGAGAGTGGC 696  
Qy 201 SerGlyProAapPheThrPheThrIleSerSerValGlnuAapLysLeuValTyr 220  
Db 697 TCTGGCGCGAATTCACCTTCAATCAATGAGTGGAGGCTGAAGACTGGACCTGAGTTAT 756  
Qy 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnLle 240  
Db 757 TTCTGTCAAGCAACATTTTCGATCTCAATCACTTCGCTGGGGAGCAAAATTTGGAAGTC 816  
Qy 241 LysalaleuGlu-----IleSerAsnSerValMetTyrPheSerSerValPro 257  
Db 817 AAAGCTTAAAGAGGGGCGGAGCGCTGGCCGGCGGACCGGACCAAGCCCTGCACCTGGCG 876  
Qy 258 ValLeuGlnLysValAsnSerThrThrLys----- 268  
Db 877 CTG-----GAGACTTTCACCCCTCATCCGCGACCGCGCGCTGGAGAACAA 921  
Qy 269 -----ProValLeuArgThrProSerProValHisProThrLysThr 282  
Db 922 CTGGAGCAGTCGGCGCTATCCGGTCGACGGCGTGGT-CGCCCTCTA---CTGGCGCGCGCG 977  
Qy 283 Ser-----GlnProGlnArgProGlnuAapCysArgProArgLysValLysGlyThr 300  
Db 978 ACTGTCATGAAACCAAGTCGACAGTGTATCCGCAACCGCCGCGACCGCCGCGAGCGG 1037  
Qy 301 GlysLeuAapPheLeuGlnuAapProLysLeuCystrLeuLeuAapGlyLleLeuPheLle 320  
Db 1038 CGGCGACCTGGGCGGAGCGATCCG----- 1061  
Qy 321 TyrGlyValLleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340  
Db 1062 -----CGAGCAGCCGCGAGCA 1076  
Qy 341 AlAlaAsnLeuGlnuAapProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlu 360  
Db 1077 GGGCCGCTGGCGCCCTGACCT-----GGCGCGCGCGCGA 1109

Qy 361 GluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGln 380  
Db 1110 GAG-----CGAGGGCTTCGTCG-----GCAGGGACCGGCA 1142  
Qy 381 ArgArgGAspAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlu 400  
Db 1143 GCACGA----- 1148  
Qy 401 AlaTyrSerGluLeuGlyThrLysGlyGluArgArgGlyLys 415  
Db 1149 -----GCCCGCGCGCCACCGCAGCGGTGAG 1178  
RESULT 8  
A50995 1862 bp DNA linear PAT 07-MAR-1997  
LOCUS Definition Sequence 36 from Patent WO9613599.  
ACCESSION A50995  
VERSION A50995.1 GI:2303794  
KEYWORDS  
SOURCE unidentified  
ORGANISM unidentified  
REFERENCE 1 (bases 1 to 1862)  
AUTHORS Wels,W. and Fominaya,J.  
TITLE NUCLEIC ACID TRANSFER SYSTEM  
JOURNAL Patent: WO 9613599-A 36 09-MAY-1996;  
WELS WINFIELD (DE)  
COMMENT Other publication AU 3926895 960523.  
FEATURES  
source  
1. .1862  
/organism="unidentified"  
/mol\_type="unassigned DNA"  
/db\_xref="taxon:32644"  
1. .1854  
/note="unamed protein product"  
/codon\_start=1  
/protein\_id="CAA03250.1"  
/translation="MDYKDDDDKKLHHHHHKLSTLEQACDLCRLKKLCKSKKPK  
AKCLKNWECSPKTRKSPPLTRALHLEVSRLERLEOLPLIPREDLMLKMSL  
ODIKALITGLPVQDNVANKDAVTRDLASVETDMPITLKHISATSSSESSNKGRDL  
TVSSSLAVGSLSCINDMDVTRDKTKTLESLKEHPINKKSESNNKVSSEKAO  
YLEERHQTALHEPLSELKTVTGNTPVPAQANVAVNAOVAVIDSTANLEKTVAA  
LSILPGIGVWGIADGAVHNTEIEIVASIALSNVAQAIPLVGLVIDIGFAYNVF  
ESIILFOVANSYNRPAYSPGVGIDKLOVLOSQSPELKKPEITVKISCKASGPF  
TNYGMNVKQAPGGGLKMMGMINSTESRPADDFKGRPDSLETSAATVLIQNNIK  
SEDMATYFCARMEYVHYGVPTWGGTITVYSSGGGSGGSGGSGSDIQLTQSHKFL  
STISGDRVSTICKASQDVNANVMTQKPGQSPKLLIYSASRYTGVPSRFTSGSGP  
DFTTISVQAEIDLAVYFCQHFPTPTFGSGTGLTETIKALDELSSERRFSA"

ORIGIN  
Alignment Scores:  
Pred. No.: 8,98e-93 Length: 1862  
Score: 1308.00 Matches: 245  
Percent Similarity: 97.64% Conservative: 3  
Best Local Similarity: 96.46% Mismatches: 6  
Query Match: 55.78% Indels: 0  
DB: 6 Gaps: 0  
US-09-596-774-7 (1-443) x A50995 (1-1862)

Qy 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 20  
Db 1090 CAGGTACACCTGCGACGACTCTGACCTGAAGCTGAGAGACAGCTCAAGATC 1149  
Qy 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40  
Db 1150 TCCGCAAGGCGCTCTGGGTATCTCTTCAAACTATGAACTGAATCGGTGAAGAGGCT 1209  
Qy 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrPheAsnThrSerThrGlyGluSerThrPhe 60  
Db 1210 CCAAGACAGGGATTAAAGTGAATGGCTGAGATTACACCTCCACTGAGAGTCAACATTT 1269

Qy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80  
Db 1270 GCTATGATCTTCANAGGACGGATTGACTTCTCTTTGAAACCTTGCCAACTGCTTAT 1329  
Qy 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100  
Db 1330 TTGCAGATCAACAACCTCAAAAGTGAAGATGATGATGATATTTCTGTGCAAGTGGAG 1389  
Qy 101 ValTyrHisSerGlyTyrValProTyrTTPGlyGlnGlyThrThrValThrValSerSerGly 120  
Db 1390 GTTATCCACGGCTACGTTCTTACTGTGGGCCAAGGACCAAGGTACCGTTTCTCTGCG 1449  
Qy 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 140  
Db 1450 GGTGGCGGTTCTGTGGCGGTGCTCCGCGGTGGCGGTTCTTACATCCAGTACCCAG 1509  
Qy 141 SerHisLysPheLeuSerThrThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160  
Db 1510 TCTCACAAAATTCCTGTCCACTTCAGTAGAGACAGGGTCAAGATCACTGCAAGGCGAGT 1569  
Qy 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180  
Db 1570 CAGATGTGTATATGCTGTGGCTGGTTCACACAGAAACAGAGACAAATCTCTAAACTT 1629  
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
Db 1630 CTGATTTACTCGGACACCTCCCGGTACACTGAGATCCCTTCTGCTTCACTGCAAGTGG 1689  
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220  
Db 1690 TCTGGCGCGGATTCTTCACTTTCACCATCAACAGTGTGACAGCTGAAGACCTGGCAGTTTAT 1749  
Qy 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240  
Db 1750 TTCTGTACGAAATTTTGTACTCTCATTCACCTTGGCTGGGCAAAATTTGAGATC 1809  
Qy 241 LysAlaLeuGluLysSerAsnSerValMetTyrPheSerSer 254  
Db 1810 AAAGCTCTAGAGATCTCTCGATGAGACAAAGTTTCAAGCC 1851  
RESULT 9  
AR268367 1862 bp DNA linear PAT 10-APR-2003  
LOCUS AR268367  
DEFINITION Sequence 36 from patent US 6498233.  
ACCESSION AR268367  
VERSION AR268367.1 GI:29698717  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 1862)  
AUTHORS Wels,W. and Fominaya,J.  
TITLE Nucleic acid transfer system  
JOURNAL Patent: US 6498233-A 36 24-DEC-2002;  
FEATURES  
source  
1. .1862  
/organism="unknown"  
/mol\_type="genomic DNA"

ORIGIN  
Alignment Scores:  
Pred. No.: 8,98e-93 Length: 1862  
Score: 1308.00 Matches: 245  
Percent Similarity: 97.64% Conservative: 3  
Best Local Similarity: 96.46% Mismatches: 6  
Query Match: 55.78% Indels: 0  
DB: 6 Gaps: 0  
US-09-596-774-7 (1-443) x AR268367 (1-1862)

Qy 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 20  
Db 1110 GAG-----CGAGGGCTTCGTCG-----GCAGGGACCGGCA 1142

[illegible]

QY	DB	Alignment Scores:	Score:	Percent Similarity:	Best Local Similarity:	Query Match:	DB:
1	1147	9.32e-93	1308.00	97.64%	96.46%	55.78%	6
21	1207	1308.00	97.64%	96.46%	55.78%	55.78%	6
41	1267	1308.00	97.64%	96.46%	55.78%	55.78%	6
61	1327	1308.00	97.64%	96.46%	55.78%	55.78%	6
81	1387	1308.00	97.64%	96.46%	55.78%	55.78%	6
101	1447	1308.00	97.64%	96.46%	55.78%	55.78%	6
121	1507	1308.00	97.64%	96.46%	55.78%	55.78%	6
141	1567	1308.00	97.64%	96.46%	55.78%	55.78%	6
161	1627	1308.00	97.64%	96.46%	55.78%	55.78%	6
181	1687	1308.00	97.64%	96.46%	55.78%	55.78%	6
201	1747	1308.00	97.64%	96.46%	55.78%	55.78%	6
221		1308.00	97.64%	96.46%	55.78%	55.78%	6

DB	1807	TTCTCTCAGCAACATTTTCGTACTCTCATTTACAGTTCGGCTCGGGACCAAAATTGGAGATC	1868
QY	241	LYSALALEUGLILLESERASMSERVALMETRYRPHESER	254
DB	1867	AAAGCTTAGAGAGATCTCTCGAGTGAGAGAAGATTTTCAGCC	1908
RESULT 11			
LOCUS	AR268366	1919 bp	linear
DEFINITION	Sequence 34 from patent US 6498233.		
ACCESSION	AR268366		
VERSION	AR268366.1	GI:29698716	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	1 (bases 1 to 1919)		
TITLE	Wels, W. and Fomlnaya, J.		
JOURNAL	Nucleic acid transfer system		
FEATURES	Patent: US 6498233-A 34 24-DEC-2002;		
SOURCE	Location/Qualifiers		
	1..1919		
	/organism="unknown"		
	/mol_type="genomic DNA"		
ALIGNMENT SCORES:			
Alignment Scores:	9,32e-93	Length:	1919
Pred. No.:	1308..00	Matches:	245
Score:	97.64%	Conservative:	3
Percent Similarity:	96.46%	Mismatches:	6
Best Local Similarity:	55.78%	Indels:	0
Query Match:	6	Gaps:	0
DB:			
US-09-596-774-7 (1-443)	x	AR268366 (1-1919)	
QY	1	GLNLEGLLEUVALGLN SERGLYPRDGLU LEU LYS LYS PRDGLYGLUTHRVALLYSILE	20
DB	1147	CAGGTACACTGCACGACACTCTGACCTGACCTGAAAGAAAGCCTGGAGAGACGTCAAGATC	1206
QY	21	SERCSYLSALASERGLYTYR-PROPHETHRASNTYRGLYMETASNTYRVALYSGLNALA	40
DB	1207	TCTCGAAGAGCCTCTGGGTATCCTTTTCAACAACATNAGAAAGAACTGGTGAACAGGCT	1266
QY	41	PROGLYINGLYLEULYSTYRMERGLYTRIPLEASNTYRTHRSERTHRGYGLUSERTHR-PHE	60
DB	1267	CCAGAGACGGGCTTAAAGTGTAGTGGCGCTGAGTTAAACCTCCACTGGAGAGTCMAACATTT	1326
QY	61	ALAASPAPPELYSGLYARGPHASPPHESELEUGLUTHRSERALASANTHRALEATYR	80
DB	1327	GCTGATGACTTCMAAGGACGGTGTGACTTCTCTTTGGAAACCTTGCCAACTGCTAT	1386
QY	81	LEUGLILLEANASNTLEULYSERGLUASPMEALATHRYRPHCYSAIARGTRPGLN	100
DB	1387	TTGCAGATCAACAACCTCAAAAGTGAAGACATGCTCATATTTCTGTGCAAGTGGAG	1446
QY	101	VALTYRTHSGLYTYRVALPRCTYRTRTPGLYGLNGLYTHRTHRVALTHRVALSERSEGLY	120
DB	1447	GTTTACACACGGCTACGTTCTTACTGTGGGGCCMAAGGACCAACGGTCCACCTTCTCTGCG	1506
QY	121	GLYGLYGLYSERGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLYGLY	140
DB	1507	GGTGGCGGTTCTGGTGGCGGTTGGCTCCGGCGGTGGCGGTTCTGACATCCAGCTGACCCAG	1566
QY	141	SERHISLYSPHELEUSERTHRSERVALGLYASPARVALSERILETHRCYLSYSAISER	160
DB	1567	TCTCACAAATTCCTGTCACACTTCAGTGAAGAGACAGGGTCAGCATCACTCCAAAGGCCAGT	1626
QY	161	GLIASPYVALTYRANALAVALLATRTYRGLNGLNLYSPROGLYGLINERPROLYSLEU	180
DB	1627	CAGGATGTGATATATGCTGTGGCTCGGTATCCMAAGAAACAGAGACAAATCTCTTAAACTT	1866
QY	181	LEULYRISERVALASERARGTYRTHRGYVALPROSERMARGPETHRNGLYSERGLY	200

Db	1667	CTGATTATCTGGGATCTCCCGGTAACCTGGATCTCTTCGCTTCACTGGAGTGGC	1746		
Oy	201	SerglyProaPhePheThrPheThrIleSerSerValGlnIleGluAspLeuAlaValTyr	220		
Db	1747	TCGGGCGGGAATTCACCTTACCATCAGCAGTGTGACGGCTGAAGACCTGGCAGTTAT	1806		
Oy	221	PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrIleLeuGlnIle	240		
Db	1807	TTCTGTGAGCAACATTTTCGTACTCCATTCACGTTCCGCTCGGGGACAAATGAGATC	1866		
Oy	241	LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSer	254		
Db	1867	AAAGCTTAGAGATCTCTCGAGTGAAGACAGATTTTCAGCC	1908		
RESULT 12	A57341	2214 bp	DNA	linear	PAT 03-MAR-1998
LOCUS	A57341	Sequence 11 from Patent EP0739584.			
DEFINITION	A57341				
ACCESSION	A57341.1	GI:3713220			
VERSION					
KEYWORDS					
SOURCE					
ORGANISM					
REFERENCE					
AUTHORS	1	Wels,W.D., Schmidt,M. and Groner,B.D.			
TITLE		Bivalent polypeptides containing at least two domains			
JOURNAL		Patent: EP 0739984-A 11 30-OCT-1996;			
		SAN TUMORFORSCHUNGS GMBH (DE)			
FEATURES					
source					
	1.	2214			
		locaton/Qualifiers			
		organism="synthetic construct"			
		/mol_type="unassigned DNA"			
		/db_xref="taxon:32630"			
		/clone="PMS238-5-TGF"			
	1.	2100			
		/note="unnamed protein product"			
		/cdon_start=1			
		/transl_table=1			
		/protein_id="CA03443.1"			
		/db_xref="GI:3713221"			
		/translation="MKKTATATATVALAGPATAQADYKDDDKLHHHHHHKQVOYOQ			
		SGELPKKPGFTAKISCAAGYPTNTGMMNVKAPQGLKMGMINITSGESTFADF			
		KGFDFSLSTANTAYLQINNLKSEDMATYFCARMEVNVGYVWQGTIVVSQGG			
		GGSGGGGGGGSDIOLTOHKPLSTGVSRAVSTICKASQDVVNAVMYQKPGQSPKL			
		LITASASRYGVPSRFSGSGSGDPFTFTTSSVOAEDLAVFCQHRPTFTFGSKPL			
		ETVALSEHHHHHHHGGSIALLTAHQACPLEFTFTRHROPGRGMELECCGYPIQRIVA			
		LYTAALPLSNQVQVQVIRNALASPGSGDGEATREPEQARLALTLAAESERFVQG			
		TGNDGAANADREKLTGTGVSHFNDPCPDSTHQPCTFHGTCAFLVQEDKPAKVCHSGTVG			
		ARCEHADLNLSTLESSAGPADSGDALERNPTGAERPLAGGGDVSFSTRGTQMWVRL			
		LOAHRLDERGVYGVVGHGTFLEMAQSIYGVGVARSDLDALIRGFTYIAGDPALVG			
		YAADQEDDAGRIIRNGALLRVYPRSSLPGFYRTSLTLAAPEAAGSEVERLIGHPLPLR			
		LDLITGTEEGEGRLLETILGWLPEAERTVIVPSALPTDPRNVGGDLDPSSIPDKQALISA			
		LPVYASQPGKPRPREDLK"			



```

Db      175 TCCTGCAAGCGCTCTGGATATCTTTCACAACTAGATGATGATGCGGTGAGCAGCT 234
Qy      41  ProGlyGlnGlyLeuLeuYrTrpMetGlyTrpIleAsnThrSerThyGlyGlnSerThrPhe 60
Db      235 CCAGGACAGGGGTTTAAAGTGATGGCTGGATTAAACCTTCACATGAGAGCAACATT 294
Qy      61  AlaAspAspPheLeuGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db      295 GCTGATGACTTCAGAGGACGGTTTACCTCTTGGAAACTCTGCCAACACTGCTCTAT 354
Qy      81  LeuGlnIleAsnAsnLeuLeuYrSerGlyAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db      355 TTGCAGATCAACAACCTCAAAAGTAGAGACATGCTACATATTTCTGGCAAGATGGAG 414
Qy      101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyTrpThrValThrValSerGly 120
Db      415 GTTAAACACGGCTAGCTTCTTACTAGGGCCCAAGGACACAGCTACCGTTCTCTGAC 474
Qy      121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db      475 GGTGGCGGTTCTGGTGGCGGTGGCTCCGGCGGTGGCGGTTCGACATCGACGTGCCAG 534
Qy      141 SerHisLeuPheLeuSerThrSerValGlyAspArgValSerIleThrCysLeuAlaSer 160
Db      535 TCTCAACAATTCCTGTCCACTTCAGTAGAGACAGAGTCAGCATCACTGCACAGGCCAGT 594
Qy      161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLeuPheProGlyGlnSerProlyLeu 180
Db      595 CAGAGTGTGTAAAGCTGTGCTGCTGATCAACAGAAACCAAGCAATCTCTTAACCTT 654
Qy      181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db      655 CTGATTTACTGGGATCTCTCCGGTACACTGAGAGTCTCTTCGCTTCACTGGCAGTGGC 714
Qy      201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaTyr 220
Db      715 TCTGGGCGGATTTCACTTTCACATCAGACAGTGTGAGCGGTGAGAGACTGGCAGTTAT 774
Qy      221 PheCysGlnGlnIleAspPheArgThrProPheThrPheGlySerGlyThrLeuGluIle 240
Db      775 TTCTGTCAAGCAATTTCTGATCTCATTCACGTTCCGGCTCGGGACAAATTTGAGATC 834
Qy      241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
Db      835 AAAGCTCTTAGACA----- 848
Qy      261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
Db      849 -----CATCATCAACATCACTAGA 869
Qy      281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGlyThr 300
Db      870 GGG---CGGCAAGCGGCGGCGGCTGAC-----CGGCAACAGGCGCTGCCACTGCCCT 920
Qy      301 GlyLeuAspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPheIle 320
Db      921 GGAGACTTTCACCCCTCATCGCCACCGCGCGCTGGGAACAACCTGAGACACTGCGGCTA 980
Qy      321 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340
Db      981 TCCGGTGACAGGGCGGTGGCTCTTACCTGGCGGCACTGCTCA----- 1026
Qy      340 PAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArgG1 360
Db      1027 -----TGGAAACAGGTCGACA-GGT- 1046
Qy      360 uGluTyrAspValLeuGluLysLysArgAlaArgAlaArgLysMetGlyGlyLysGlnG1 380
Db      1047 -----CATCGCAAGCGCTCGGCCACCCCG 1072
Qy      380 nArgArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400
Db      1073 GCAGCGGCGGAGACTCTGGGCGAAGGATCCGGACAGCCGAGACAGGCCCGCTCTGGCC 1132

```

```

Qy      400 LuAlaTyrSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp 418
Db      1133 TGACCTGCGCGCGCGCGGAGAGCGACGCGCTTCCTCGGACAGGACCGGCAACGAC 1188

RESULT 13
LOCUS      A57343          2793 bp    DNA          linear    PAT 03-MAR-1998
DEFINITION Sequence 13 from Patent EP0739984.
ACCESSION  A57343
VERSION     A57343.1   GI:3713222
KEYWORDS   synthetic construct
SOURCE      synthetic construct
ORGANISM    other sequences; artificial sequences.
REFERENCE   1
AUTHORS     Wels,W.D., Schmidt,M. and Groner,B.D.
TITLE       Bivalent polypeptides containing at least two domains
JOURNAL     Patent: EP 0739984-A 13 30-OCT-1996;
            SAN TUMORFORSCHUNGS GMBH (DE)
FEATURES
   source
       1..2793
           /organism="synthetic construct"
           /mol_type="unassigned DNA"
           /db_xref="taxon:32630"
           /clone="PMS238-5-225"
       1..2679
           /note="unnamed protein product"
           /codon_start=1
           /transl_table=1
           /protein_id="CAA03444.1"
           /db_xref="gi:3713223"
           /translation="MKKTAIALVALAGFAVDAQDYKDDDKLHHHHHKLQVLOO
SPELKKRGEVVKISCAKSGPTPTNGMAYKQAPGGLKRMGINSTGSTRADPF
KGRDFSLKETSANTAYLIQINLKSEDMATTFCAWEVYHGVPTVWGDTTVSGG
GSGGGGSGGGSQIQLTQSHKFLSTSVGQEDLAIVYQDVNVAWYQKPGQPKL
LIYASASRYGTPVSRFTGSGSPDTFTLISVQMEDLAIVYQCOHFRPTFGSGTKL
EIKALEHHHHHHLEGSIALTAHOACHLPLETTRHROPGMQLBOCGVPVORLVA
LYLAARLSMNOVDVIRNALAPSGGDGLGAIIEOPBOARLALTAAASERFVROG
TGNDACGANDEKLOVLOESGPGVLOPSGLSTCTTCSGFSFLTNGVHYRSPK
GLEWLGVYWGSGNTDYNTPFTSRLSINKDNKSOVFFKMSLQSDNATLYCANALTY
YDFEFAVWGQITVTVSSGGSGGSGGSGGSGGSDTFLSINSVESCSRA
SOSIGTINHWYQKRTNGSPRLIKYASESISGIPRSFGSGSDTFLSINSVESDI
ADYVCOONMMPTFRGAGTKLEIKALDPRPSAGPADSGDALERNYPGAEFLD
GGVSRSTRGQNTVRLIOAHQOLERGVFPVGYCTGLEAOSIVFGVTRASOD
LDATWRGPTVLDGPLALVGYAODEPARGKIRIGALLRYTPVPSLSLPGFRISLTLA
APENAGVERLIGHPLRLDALTGPBEEGRLETLIGWPLAERTVIVPSAIPDPN
VGDLDPSSIPDKQKAISALPDVYASQPKPRBDLK"

ORIGIN
Alignment Scores:
Pred. No.:      1,49e-92      Length:      2793
Score:          1308.00      Matches:      275
Percent Similarity: 71.19%      Conservative: 24
Best Local Similarity: 65.48%      Mismatches: 59
Query Match:    55.78%      Indels:      64
Db:             6          Gaps:      5

US-09-596-774-7 (1-443) x A57343 (1-2793)
Qy      1  GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 20
Db      115 CAGGTACACTGCAGCACTCTGGACCTGACATGAAGAAGCTTGAGAGACAGCTCAAGTTC 174
Qy      21  SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db      175 TCCTGCAAGCGCTCTGGATATCTTTCACAAACTATGGAATGAACTGGGTGAAACAGGCT 234
Qy      41  ProGlyGlnGlyLeuLeuYrTrpMetGlyTrpIleAsnThrSerThyGlyGlnSerThrPhe 60
Db      235 CCAGGACAGGGGTTTAAAGTGATGGCTGGATTAAACCTTCACATGAGAGCAACATT 294
Qy      61  AlaAspAspPheLeuGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80

```

```

Db      ||| 295 GCTGATGACTTCAGAGGAGCGTTTACTTCTTGGAACCTCGCAAGCTCCTAT 354
Qy      ||| 81  LeuGlnIleAsnAsnLeuLeuSerGluAspMetAlaThrTrpPheCysAlaArgTrpGlu 100
Db      ||| 335 TTGCGATCAACAACCTCAAAAGTAAAGACATGCGTACATATTTCTGGCAAGAGGAG 414
Qy      ||| 101 ValTrpHisGlyTrpValProTyrTrpGlnGlnGlyThrThrValThrValSerSerGly 120
Db      ||| 415 GTTACACACCGGTACTGTTCTTACTGGGGCCAGAGGACACCGGTACCGTTCTCTGGC 474
Qy      ||| 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
Db      ||| 475 GGTGGCGGTTCTGGGCGGTGGCTCCGCGGTGGCGGTCTGCATCATCACTGACCGAC 534
Qy      ||| 141 SerHisLeuPheLeuSerThrSerValGlyAspArgValSerIleThrCysLeuAlaSer 160
Db      ||| 535 TCTCACAAAATTCCTGTCCTCACTCAAGAGAGAGGAGTCAAGATCACTGCAAGGCGAGT 594
Qy      ||| 161 GlnAspValTrpAsnAlaValAlaTrpTyrGlnGlnGlnGlnGlnGlnGlnGlnGln 180
Db      ||| 595 CAGGATGCTGTAATAGTCTGTTGCGGTATCAACAGAAACAGAGCATCTCTAAACTT 654
Qy      ||| 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db      ||| 655 CTGATTACTCGGATCTCCCGGTACACTGAGAGTCTCTCTGCTTCACTGCGAGTGGC 714
Qy      ||| 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyr 220
Db      ||| 715 TCTGGGCGGATTCATCTTCACTCAACATCAAGAGTGTGAGGCTGAAAGCTGGCAGTTAT 774
Qy      ||| 221 PheCysGlnGlnIleAspArgTrpProPheThrPheGlySerGlyThrTrpLeuGlnIle 240
Db      ||| 775 TTCTGTCAGCAACATTTCTGACTCCATTCAGTTCGGCTCGGGGCAAAATGGAGATC 834
Qy      ||| 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
Db      ||| 835 AAACCTCTAGAGCA----- 848
Qy      ||| 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
Db      ||| 849 -----CATCATCAACCTCACTCAAGA 869
Qy      ||| 281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThr 300
Db      ||| 870 GGG---CGGCGAGCTGGCGCGCTGAC---CGCGCACCAAGCGCTGCCACCTGCGCGCT 920
Qy      ||| 301 GlyLeuAspPheLeuGlnAspProLysLeuGlyTyrIleLeuAspGlyIleLeuPheIle 320
Db      ||| 921 GGAAGACTTTCACCCGTCATCGCCAGCGCGCTGGGAAACACTGAGAGCATGCGGCTA 980
Qy      ||| 321 TyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340
Db      ||| 981 TCGGTGAGAGCGGTGGTGGCTCTTACTGCGCGCGCACTGCA----- 1026
Qy      ||| 340 PAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGly 360
Db      ||| 1027 -----TGGAAACAGGTGCAACA-GGT----- 1046
Qy      ||| 360 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnMetCysLysGlnGln 380
Db      ||| 1047 -----GATCCGCAACGCCCTGCGCAAGCGCG 1072
Qy      ||| 380 ArgArgArgAsnPro-GlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400
Db      ||| 1073 GCAAGCGCGGAGACTCGGGGAAAGCATCGCGAGCAGCGCGAGCAGCGCGCTGCGCCC 1132
Qy      ||| 400 LysAlaTyrSerGlnIleGlyThrLysGlyLysArgArgGlyLysGlyHisAsp 418
Db      ||| 1133 TGACCTTGCGCGCGCGAGAGCGGCTTCTGCGCGAGGCAACCGCAACGAC 1188

```

```

LOCUS      A57351                2793 bp    DNA             linear      PAT 03-MAR-1998
DEFINITION Sequence 21 from Patent EP0739984.
ACCESSION  A57351
VERSION     A57351.1   GI:3713230
KEYWORDS
SOURCE
  ORGANISM
    1. synthetic construct
    2. synthetic construct
    3. other sequences; artificial sequences.
REFERENCE
  1. Wels,W.D., Schmidt,M. and Groner,B.D.
  Bivalent polypeptides containing at least two domains
  Patent: EP 0739984-A 21 30-OCT-1996;
  SAN TUMORFORSCHUNGS GMBH (DE)
  Location/Qualifiers
    1..2793
      /organism="synthetic construct"
      /mol_type="unassigned DNA"
      /db_xref="taxon:32630"
      /clone="PMS238-5-5"
      /note="MOUSE; P.AERUGINOSA; SYNTHETIC"
      1..2679
        /note="unnamed protein product"
        /codon_start=1
        /transl_table=1
        /protein_id="CAA03448.1"
        /db_xref="GI:3713231"
        /translation="MKKTAIALAVALAGFATVAQADYKDDDKLHHHHHKLQVOLO
        SGRBLKKPGETVYKISKASGYFTTYNGMNVKQAPQGLKMKGHINTSGESTFADP
        KGRFDSLSTANTAYLQINLKSEMATYFCARMEVGVYVWQGVSTTVVSSGG
        GSGGGSGGSGSDIQLTQSHKELSTVGSRVSTICASQDVNAAVMYQKQSGSL
        LIYSASRYTGVSPRFTGSGSPDPFTTISVOAEDLAVYFCQHRFRFTFGSGTKL
        ELIYALHHHHHHLEGGSLAALTAHQHKLPLEFTFHRQPRGMEQLQCGSPVORLVA
        LVYLAARLSMNVOOVIRNMLASPGSGDGEALIREQEPARLALTAAASRRFVROG-
        TGNDEAGANANDEKXLOVOLOSGPRLKRGFTVXISKSGYFPTNYGMNVKQAPQ
        GLKMKGHINTSGESTFADPFGKRPFSLETANTRYLDQINLKSMDATTYCARNEV
        YHEVYVWQGVSTTVVSSGGSGSGGSDIQLTQSHKFLSTVSGDRVSITCKA
        SQDVNAAVMYQKQSGSPKLIYSASRYTGVSPRFTGSGSPDPFTTISVOAEDL
        AVYFCQHRFRFTFGSGTKLEIKALEDRVPSGAPDAGCALLEARNPTAEFLGD
        GGVSPSTRTQMTVERLLOAHROLEBERGVYVGHTEFLAAGSIVAGGVARNSD
        LDAIWMGFYIADGPLALVGYADQDEPDARGIRRNALILRVYVPRSLIPFVYTSLLTA
        APEAAQEVRIIGHPLPIPLDAITGPEBEGSLFTILGMLAEIRVIVIPSLPTPBRN
        VGSDLPSSIPDKXONISALPDYASQPKRPREDLK"
ORIGIN
Alignment Scores:
Pred. No.:      1,49e-92      Length:      2793
Score:          1308.00      Matches:      275
Percent Similarity: 71.198      Conservative: 24
Best Local Similarity: 65.488      Mismatches: 59
DB:              55.784      Indels:      64
              Gaps:      5
US-09-596-774-7 (1-443) x A57351 (1-2793)
Qy      ||| 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysLysPheProGlyGluThrValLysIle 20
Db      ||| 115 CAGGTCAACCTGAGAGAGCTGAGACTGAAGTGAAGAACCTGAGAGACAGTCAAGATC 174
Qy      ||| 21 SerCysLeuAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db      ||| 175 TCTGCAAGGCGCTCGGGATCTTTCACAAACTATGAAATACCTGGGTGAAGAGCGCT 234
Qy      ||| 41 ProGlyGlnGlyLeuLysTrpMetGlyTrpIleAsnThrSerThrGlyGlnSerThrPhe 60
Db      ||| 235 CCAAGACAGAGGTTTAAAGTGAATGGGCTGATTAACACTCACTCACTGAGAGTCAACATTT 294
Qy      ||| 61 AlaAspAspPheLysGlyLysArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80
Db      ||| 295 GCTGATGACTTCAAGGAGCGGTTTGAATCTTCTTGGAACCTCTGCAACACTGCTAT 354
Qy      ||| 81 LeuGlnIleAsnAsnLeuLeuSerGluAspMetAlaThrTrpPheCysAlaArgTrpGlu 100
Db      ||| 355 TTGCGATCAACAACCTCAAAAGTAAAGACATGCGTACATATTTCTGTCGAAGATGGAG 414

```

```

Qy 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSergly 120
Db 415 GTTATACACAGCGGTACGTTCTTACTGCGGCCAAGGACACAGCTTCTCTGCGC 474
Qy 121 GlyGlyGlySerglyGlyGlySerglyGlyGlySerglyGlyGlySergly 140
Db 475 GGTGCGGCTTCTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCGGTGCG 534
Qy 141 SerHisSerPheLeuSerThrSerValGlyAspArgValSerThrCysValValSer 160
Db 535 TCTCACAAAATCTCTTCCACTTCAGTAGAGACAGAGGTGATGATCACTGCAAGGCGAGT 594
Qy 161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db 595 CAGAGTGTATATAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 654
Qy 181 LeuIleTyrSerAlaSerSerSerglyTyrThrGlyValProSeraArgPheThrGlySergly 200
Db 655 CTGATTTACTCGGCATCTCTCCCGGTACACTGGAGTCCCTTCTGCTCACTGCGAGTGGC 714
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlyAspLeuAlaValTyr 220
Db 715 TCTGGGCGGATTTCACTTTCACATCAGCAGTGTGCAAGGCTGAGACCTGCGAGTTAT 774
Qy 221 PheCysGlnGlnHisPheArgThrPhePheThrPheGlySerglyThrLysLeuGlnIle 240
Db 775 TTCTGTCAACAATTTCTGATCTCATTCATTCAGTTTCTGCGGACAAATTTGAGATC 834
Qy 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 260
Db 835 AAAGCTTAGAGCA----- 848
Qy 261 LysValAlaSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
Db 849 ----- 869
Qy 281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThr 300
Db 870 GGG---CGGCAAGCTGCGCGCGCGCTGAC-----CGGCAACAGCGCTGCGACCTGCGCT 920
Qy 301 GlyLeuAspPheLeuGlnAspProLysLeuGlyTyrLeuLeuAspGlyIleLeuPheIle 320
Db 921 GGAAGCTTTCACCGCTCATCGCCAGCGCGCGCTGGGAAACAATCGAGACAGTGGCGCTA 980
Qy 321 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerAlaGlyThr 340
Db 981 TCCGGTGCAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1026
Qy 340 AlaAlaAlaSerLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGly 360
Db 1027 -----TGGAAACAGGCTGAGCA-GGT----- 1046
Qy 360 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnLysGlyLysGlnI 380
Db 1047 -----GATCGGCAAGCGCTGCGACCGCCG 1072
Qy 380 ArgArgArgArgAsnPro-GlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400
Db 1073 GCACGCGCGCGAGCCTGCGGCAAGCATCGCGAGACACCGGAGAGCGCGCTGCGGCC 1132
Qy 400 LysAlaTyrSerGlnIleGlyThrLysGlyGlyLysArgArgArgGlyLysLysHisAsp 418
Db 1133 TGACCTGCGCGCGCGCGAGAGCGGCTTCTGCTGCGAGGACCGGCAACGAC 1188

```

```

RESULT 15
LOCUS A57347 3177 bp DNA linear PAT 03-MAR-1998
DEFINITION Sequence 17 from Patent EP0739884.
ACCESSION A57347 GI:3713226
VERSION A57347.1
KEYWORDS
synthetic construct

```

## ORGANISM

synthetic construct

other sequences; artificial sequences.

## REFERENCE

## AUTHORS

Wels, W.D., Schmidt, M. and Groner, B.D.

## TITLE

## JOURNAL

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

## CDS

## FEATURES

## SOURCE

```

US-09-596-774-7 (1-443) x A57347 (1-3177)
Qy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysLysProGlyGlyThrValLysIle 20
Db 115 CAGGACACTGCGACGCTGCGACCTGAGCTGAGAGAGCGCTGAGAGACAGCTCAAGATC 174
Qy 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
Db 175 TCTGCAAGGCTCTGGGTATCTTTCACAAACATATGGAATGAACTGGGTGAAGAGCT 234
Qy 41 ProGlyGlnGlyLeuLysLysTrpMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60
Db 235 CCAGAGACAGGCTTAAAGTGGATGGCGGATTAACCTTCACCTGAGAGATCAACATTT 294
Qy 61 AlaAspAspPheLysGlyArgArgPheSerLeuGlnTrpSerAlaAsnThrAlaLysTrp 80
Db 295 GCTATATGCTTCAAGGAGAGCGGTGACTCTCTTGGAAACCTTCCGCAACAGCTGCTAT 354
Qy 81 LeuGlnIleAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db 355 TTGCAAGTCAACACCTCAAAATGGAAGATGCTCATATTTCTGTGCAAGATGGAG 414
Qy 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSergly 120
Db 415 GTTATACACAGCGGTACGTTCTTACTGCGGCCAAGGACACAGCTTCTCTGCGC 474

```

## ORIGIN

## Alignment Scores:

```

Pred. No.: 1,75e-92 Length: 3177
Score: 1308.00 Matches: 275
Percent Similarity: 71.19% Conservative: 24
Best Local Similarity: 65.48% Mismatches: 59
Query Match: 55.78% Indels: 64
Db: 6 Gaps: 5

```

QY 121 GLYGLYSERGLYGLYGLYSERGLYGLYSERASPILLEGLNLEUTHRGIN 140  
DB 475 GGTGGCGGTTCTGTGGCGGTGGCTCCGGCGGTGGCTTGACATCCAGTACCAG 534  
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLeuAlaSer 160  
DB 535 TCTGACAAATTCCTGTCCATTCAGTACAGAGACAGGCTCAGCATCAGTCCAGGCCAGT 594  
QY 161 GLNAspValTyrAsnAlaValAlaTPrTyrGlnGlnLysProGlyGlnSerProLysLeu 180  
DB 595 CAGGATGTGTAAATGCTGTGTGCTGTATCAACAGAAACAGGCAATCTCTAAACTT 654  
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
DB 655 CTGATTTACTGGCATCTCCCGGTACCTGAGTCCCTCTCTGCTTCACTGGCAGTGGC 714  
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220  
DB 715 TCTGGGCGGATTTCACTTACCATCAGAGTGTGCAAGGCTGAAGACTGGCAGTTTAT 774  
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240  
DB 775 TTCTGTGACACATTTTCGTATCCATTCAGTTCCGCTGGGACAAATTTGAGATC 834  
QY 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260  
DB 835 AAAGCTCTAGACA----- 848  
QY 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280  
DB 849 -----CCATCATCACCATCAGCTAGA 869  
QY 281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgLysSerValLysGlyThr 300  
DB 870 GGG---CGGCGCTTGGCGCGCTGAC-----CGCGCACCGAGCTGCGCACTGCCGCT 920  
QY 301 GlyLeuAspPheLeuGlnAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPheIle 320  
DB 921 GGAGACTTTTACCGGTATCGCCAGCGCGCGGCTGGGAAACACTGAGACAGTCCGCTA 980  
QY 321 TyrGlyValIle-IleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThr 340  
DB 981 TCCGCTCAGCGGCTGTGCTGCTCTTACTGCGCGCGCACTGTCA----- 1026  
QY 340 rAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGly 360  
DB 1027 -----TGGAACCGAGTGCACCA-GGT----- 1046  
QY 360 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGlnMetGlyGlyLysGlnI 380  
DB 1047 -----GATCCGCAACGCCCTGCGCCAGCCCCG 1072  
QY 380 nArgArgArgAsnPro-GlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400  
DB 1073 GCACGGCGCGGAGCTGGCGAAACGATCCGCGAGCAGCCGAGCAGGCCCTGTGCCCC 1132  
QY 400 IuAlaTyrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp 418  
DB 1133 TGACCTGCGCGCGCGGAGCGAGCGCTTCTCCGCGAGGCGACCGGCAACGAC 1188

Search completed: April 20, 2005, 14:06:52  
Job time : 5971.62 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 20, 2005, 07:10:49 ; Search time 706.644 Seconds  
(without alignments)  
3711.128 Million cell updates/sec

Title: US-09-596-774-7

Perfect score: 2345

Sequence: 1 QIQVQSGPELKKRGETVKI.....LSTAKTQYDALHQTIAPR 443

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODEL=frame\_plus\_p2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool\_h/US05956774/runat\_20042005\_075441\_12501/app.query.fasta\_1.1230  
-DB=N Geneseq -QFMT=fastlap -SUFFIX=apir20.rng -MINMATCH=0.1 -LOOPCL=0  
-LOOEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=Dlosum62 -TRANS=human40.cdi  
-LIST=45 -LOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pro -NOR=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000  
-USER=US05956774 @CGN\_1\_1\_653 @runat\_20042005\_075441\_12501 -NCPU=6 -ICPU=3  
-NO MMAP -LARGOQUERY -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N\_Geneseq\_16Dec04: \*  
1: geneseqn1980s: \*  
2: geneseqn1990s: \*  
3: geneseqn2000s: \*  
4: geneseqn2001as: \*  
5: geneseqn2001bs: \*  
6: geneseqn2002as: \*  
7: geneseqn2002bs: \*  
8: geneseqn2003as: \*  
9: geneseqn2003bs: \*  
10: geneseqn2003cs: \*  
11: geneseqn2003ds: \*  
12: geneseqn2004as: \*  
13: geneseqn2004bs: \*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2245	100.0	1479	2	AAT05783 Leader-8c
2	1636.5	69.8	1452	2	AAV73324 Chimeric
3	1501.5	64.0	1284	2	AAV77139 Single ch
4	1501.5	64.0	3058	2	AAV77137
5	1312.5	56.0	2070	2	AAT42036 Plasmid p

6	1312.5	56.0	2730	2	AAT42044
7	1309.5	55.8	2011	2	AAQ28262
8	1308	55.8	1862	2	AAT29413
9	1308	55.8	1919	2	AAT29412
10	1308	55.8	2214	2	AAT42038
11	1308	55.8	2793	2	AAT42039
12	1308	55.8	2793	2	AAT42043
13	1308	55.8	3177	2	AAT42041
14	1305	55.7	1692	2	AAT29409
15	1303	55.6	2233	2	AAQ28257
16	1302	55.5	794	2	AAT42034
17	1302	55.5	2718	2	AAT42042
18	1302	55.5	2793	2	AAT42040
19	1285.5	54.8	748	2	AAT05781
20	1276.5	54.4	748	2	AAQ28256
21	1249	53.3	1920	10	ADH34695
22	1249	53.3	6834	10	ADH34699
23	1015.5	43.3	7654	8	ABX16565
24	1015.5	43.3	7654	10	ABX13168
25	1013	43.2	726	4	AAF30723
26	1011.5	43.1	1422	2	AAT90509
27	1011.5	43.1	1545	2	AAT90510
28	994	42.4	771	4	AAQ8584
29	994	42.4	771	9	AAT43523
30	981.5	41.9	909	2	AAT36878
31	981.5	41.9	909	2	AAT91837
32	981.5	41.9	909	2	AAT78879
33	981.5	41.9	909	2	AAV22598
34	981.5	41.9	909	2	AAV21796
35	981.5	41.9	909	2	AAV63397
36	981.5	41.9	909	10	AAQ61483
37	975.5	41.6	987	10	ADG32297
38	975.5	41.6	987	10	ADG32348
39	972.5	41.5	2001	2	AAQ45088
40	962.5	41.0	1956	2	AAT90512
41	962.5	41.0	2079	2	AAT90513
42	960.5	41.0	777	10	ADG32299
43	960.5	41.0	987	10	ADG32350
44	930	39.7	1350	2	AAT97844
45	928.5	39.6	800	3	AAA08034

# ALIGNMENTS

RESULT 1					
ID	AAT05783	standard; cDNA; 1479 BP.			
XX					
AC	AAT05783;				
XX					
DT	16-MAR-1996	(first entry)			
XX					
DE	Leader-scFv(FRPs):lyt-2 hinge;zeta cDNA.				
XX					
KW	Single chain antibody; scFv; antibody engineering; antitumour;				
KW	tumour antigen binding; T-cell receptor; cytotoxic T-lymphocyte;				
KW	monoclonal antibody; ecdB-2; cancer; cell targeting; gene therapy;				
KW	adoptive immunotherapy; ss.				
XX					
OS	Synthetic.				
XX					
FT	Key	Location/Qualifiers			
FT	CDS	40..1426			
FT		/*tag= a			
FT		/product= "Leader-scFv(FRPs):lyt-2 hinge;zeta"			
XX					
XX					
PD	09-NOV-1995.				
XX					
PF	20-APR-1995;	95WO-EP001494.			
XX					
PR	02-MAY-1994;	94EP-00810244.			

XX (CIBA ) CIBA GEIGY AG.  
 XX  
 XX Groner B, Moritz D;  
 XX  
 XX WPI, 1995-393085/50.  
 DR P-PSDB; AAR85508.  
 XX

PT New bifunctional proteins for use in killing tumour cells - contg. a  
 PT tumour antigen binding domain, a hinge region and a zeta chain derived  
 PT from a T-cell antigen receptor.  
 XX

PS Example 1; Page 29-31; 46pp; English.

CC cDNA (AAT05783) codes for a bifunctional protein (AAR85505) consisting of  
 CC single chain antibody scFv(FRP)5 directed against the tumour erbB-2  
 CC antigen, a hinge region, and a functional zeta chain obtd. from a T-cell  
 CC receptor. The cDNA is expressed in host cells, esp. cytotoxic T-  
 CC lymphocytes, providing them with a defined tumour cell specificity  
 CC enabling targeting to defined tumour cells and MHC-unrestricted and MHC-  
 CC independent tumour destruction in vitro or in vivo  
 XX

SQ Sequence 1479 BP; 369 A; 395 C; 394 G; 321 T; 0 U; 0 Other;

#### Alignment Scores:

Pred. No.: 3e-177 Length: 1479  
 Score: 2345.00 Matches: 443  
 Percent Similarity: 100.00% Conservative: 0  
 Best Local Similarity: 100.00% Mismatches: 0  
 Query Match: 100.00% Indels: 0  
 Gaps: 0

US-09-596-774-7 (1-443) x AAT05783 (1-1479)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysPProGlyGlnThrValLysIle 20  
 DB 94 CAGATCCAGTGGTGTGACCTGTGACCTGAGCTGAAGAGCCTGAGAGACAGTCAAGTTC 153  
 QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40  
 DB 154 TCCGTGCAAGGCTCTGGGTATCCTTTCACAAACTATGAGATGAATCTGGGTAAAGCAGCT 213  
 QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrIleAsnThrSerThrGlyGlnSerThrPhe 60  
 DB 214 CCAGAGACAGGGTTAAAGTGTGAGGTGGATTAACCTCCACCTGGAGATCAACATTT 273  
 QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaSerThrAlaTyr 80  
 DB 274 GCTATGACTTCAAGGAGGAGCGGTGACTTCTTTGGAAACCTGTGCCAACACTGCCTAT 333  
 QY 81 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTrpGlu 100  
 DB 334 TTGCAAGATCAACAACTTCAAAAGAGAGACATGCTACATATTTCTGTGCAAGATGGAG 393  
 QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120  
 DB 394 GTTACACACGGCTACGTTCTTACTGGGGCCAAAGGACACAGGTTCACGTTTCTCTGGC 453  
 QY 121 GlyGlyLysSerGlyGlyGlyLysSerGlyGlyGlyLysSerAspIleGlnLeuThrGln 140  
 DB 454 GGTGGCGGTCTGTGTGGCGGTGCTCCGGCGGTGGCTTCTGCAATCCAGCTGAGCCAG 513  
 QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160  
 DB 514 TCTACAAATTCCTGTCTCACTTCACTGAGAGACAGGGTCAACATCCTGCAAGGCCAGT 573  
 QY 161 GlnAspValTyrAsnAlaValAlaTyrGlnGlnLysPProGlyGlnSerProLysLeu 180  
 DB 574 CAGGATGTGATATGCTGTGCTGTGATCAACAGAAACAGAGCAATCTCCCTAACTT 633  
 QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
 DB 634 CTGATTTACTGGGATCCTCCGGGTACAGTGAAGTCCCTTCTGCTTCACTGGGAGTGGC 693

QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnIleAspLeuAlaValTyr 220  
 DB 694 TCTGGCCGGATTTCACTTTCACCATCAGAGGTGAGCTGAAGCTGGGAGCTTTAT 753  
 QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240  
 DB 754 TTTCTGAGCAACAATTTTCTGATCTCCATTCACGTTGGCTCGGGGACAAATTTGAGATC 813  
 QY 241 LysAlaLeuGlnLysSerAsnSerValMetTyrPheSerSerValProValLeuGln 260  
 DB 814 AAGCTCTAGAGATCAGCAACTCGGTGATGATCTTCACTGCTGCGAGCTTTCAG 873  
 QY 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280  
 DB 874 AAGTGAACCTCTACTACTACCAAGCCAGTGTGCTGAACTCTCCTACCTGTGACCTTACC 933  
 QY 281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValLysGlyThr 300  
 DB 934 GGGACATCTAGCCCCAGAGACCAAGAAATTGTCGGCCCGGTGGCTCAGTGAAGGGGACC 993  
 QY 301 GlyLeuAspPheLeuGlnAspProLysLeuCysTyrIleLeuAspGlyIleLeuPheIle 320  
 DB 994 GGATTTGACTTCTAGAGATCCCAAACTCTGCTACTGTGATGGAATCCTTCTATC 1053  
 QY 321 TyrGlyValIleIleThrAlaLeuTyrIleuArgAlaLysPheSerArgSerAlaGluThr 340  
 DB 1054 TACGAGTCACTATCAGAGCCCTGTACTGTGAGGACAAATTCAGCAGAGTGCAGAGACT 1113  
 QY 341 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlu 360  
 DB 1114 GCTGCCAAGCTGAGAGCCCAACCAAGCTCTCAATGAGCTCAATTAAGGGGCAAGAG 1173  
 QY 361 GluTyrAspValLeuGlnLysLysArgAlaArgAspProGluMetGlyLysGlnGln 380  
 DB 1174 GAATATGACCTCTTGGAGAGAGAGCGGCTCGGATCCAGATGGAGGACAAACAGCAG 1233  
 QY 381 ArgArgArgAspProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlu 400  
 DB 1234 AGAGAGAGAAACCCCAAGAGGCGTATTCATGCTCAGCAAGAAAGCAAGATGGCAGAA 1293  
 QY 401 AlaTyrSerGluIleGlyThrLysGlyGluArgArgArgGlyLysGlyHisAspGlyLeu 420  
 DB 1294 GCTTACAGTGAATCGGACAAAGGCGAGAGCGGAGGAGGAGGACAGATGGCTTT 1353  
 QY 421 TyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThrLeu 440  
 DB 1354 TACCAAGGCTCTCAGACTGCACCAAGAGCACCTATGATGCCCTGATATGACAGACCTTG 1413  
 QY 441 AlaProArg 443  
 DB 1414 GCCCTTCGC 1422

RESULT 2  
 AAV73324 standard; DNA; 1452 BP.  
 ID AAV73324  
 XX  
 AC AAV73324;  
 XX  
 DT 26-FEB-1999 (first entry)  
 XX  
 DE Chimeric CD19/CTCR DNA.  
 XX  
 KW Single chain; antibody; B-cell marker; CD19; scFvCD19; gene therapy;  
 KW retroviral vector; tumour-associated antigen; cancer; immunoglobulin;  
 KW CD8 alpha chain; hinge region; theta chain; transmembrane domain;  
 KW intracellular domain; T-cell receptor; CD3 complex; ss.  
 XX  
 OS Homo sapiens.  
 OS Synthetic.  
 XX  
 PN DE19720152-A1.  
 XX

PD 05-NOV-1998.  
PE 02-MAY-1997; 97DE-01020152.  
PR 02-MAY-1997; 97DE-01020152.  
PA (DELB-) DELBRUECK CENT MOLEKULARE MEDIZIN MAX.  
XX Schroeder A, Schwenkenbecher J, Bargou R;  
XX WPI; 1998-584686/50.  
XX P-PSDB; AAW82345.  
XX Retroviral vector system for gene therapy - by expression of therapeutic  
PT gene in T cells.  
PS  
PS Claim 9; Page 6; 10pp; German.  
XX  
XX This sequence encodes a novel chimeric CD19/CTCR protein which is  
CC composed of an immunoglobulin leader sequence, an anti-CD19 single chain  
CC antibody fragment (ScFvCD19), a hinge region from the CD8 alpha-chain and  
CC the transmembrane and intracellular domain of the T-cell receptor-CD3  
CC complex theta chain. This sequence is used in a retroviral vector system  
CC for gene transfer and expression in primary human T cells which involves  
CC a retroviral vector containing cDNA corresponding to a therapeutic gene,  
CC a promoter selected from the SV40 immediate early promoter or the human  
CC IL-2 or MHC I promoter and a selectable marker, and a packaging cell line  
CC (i.e. a modified retroviral packaging line for enhanced transfer into  
CC primary human T lymphocytes). This vector system in which the cDNA  
CC encodes a single-chain antibody (scFv) specific for a tumour-associated  
CC antigen can be used for gene therapy of cancer, especially by isolating T  
CC cells from the patient's blood, stimulating the T cells in vitro, and  
CC introducing the retroviral vector into the T cells and returning the T  
CC cells to the patient, preferably by intravenous injection

Alignment Scores:	
Pred. No.:	7, 77e-121
Score:	1636.50
Percent Similarity:	78.45%
Best Local Similarity:	70.26%
Query Match:	69.79%
EB:	2
length:	1452
Matches:	326
Conservative:	38
Mismatches:	79
Indels:	21
Gaps:	5

US-09-596-774-7 (1-443) x AAV73324 (1-1452)

QY 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysLysProGlyGlnThrValIle 20  
Db 58 GAGTACACATGCAGAGACTCAGGGGCTGAGCTGATGAGGCTGGCTCCAGTGAAGATT 117  
QY 21 SerCysValSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIleGlnAla 40  
Db 118 TCCGTCAGAGGCTCTCGGCTATCATTCAGTACTACTGATGAATCAGGGTGAAGCAGAG 177  
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrPheIleAsnThrSerThrGlyGlnSerThrPhe 60  
Db 178 CCTGACACGGGCTTGAGTGAATTGGACAGATTGGCTCGAGATGGATGATTAATCACTAC 237  
QY 61 AlaAspAspPheLysGlyValArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80  
Db 238 AATGGAAGTTCMAAGGTTAAACCCACTCTGACTGCAGAGCAATCTCCAGCACAGCTTAC 297  
QY 81 LeuGlnIleAsnAsnLeuLysSerGlnAspMetAlaThrTyrPheCysAlaArgTrpGln 100  
Db 298 ATGCACACTCAGACGCTAGCATCTGAGGACTCTGGCGGTATTTCTGTGCAGAACGGGAG 357  
QY 101 Val-----TyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrVal 115  
Db 358 ACTACGACGGTAGCGCTTATTACTATGCTATGATGACTACTGGGGGTCAAGAACTCATGC 417  
QY 116 ThrValSerSerGlyGlyGlyGlySer-----GlyGlyGlyGlySerGlyGlyGlyGly 133  
Db 116 TTTTATTTT

Db	418	ACCGCTCTCCAGCAAAAAGCAGACCCAGCTTGAGAAAGGTGAATTTTCAGAAAGCAGC	477
Qy	134	Se7Ap11egInLeuThrInGlnSerHisLysPheLeuSerThrSerValGlyAspArgVal	153
Db	478	GTAATATCTTGATGACCCCAATCTCCAGCTTCTTTGGCTGGGCTTGAAGGACAGGGCC	537
Qy	154	Se11eThrCysLysAlaSerGlnAspVal---TyrAsnAla-----ValAlaThr	169
Db	538	ACCAATCTCTGCAAGGCCAGCCAAAGGTGTATTATGATGAGTATAGTTATTGAATCTGG	597
Qy	170	TyrGlnGlnLysProGlnGlnSerProLysLeuLeuLeuTyrSerAlaSerSerArgTyr	189
Db	598	TACCAACAGATCTCAGAGCAGCCACCAGAACTCTCATCTATGATGATCAATCTAGTT	657
Qy	190	ThrGlyValProSerArgPheThrGlySerGlySerGlyProAspPheThrPheThrIle	209
Db	658	TCTGGGATCCCAACCGAGTTTAAAGTGGCAGTGGTCTGGGACAGACTTCACCTCAACATTC	717
Qy	210	SerSerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGlnHisPheArgThrPro	229
Db	718	CATCCTGGAGAAAGTGGATGCTGCAACCTATCACTGTCAAGAAAGTACTGAGGATCCG	777
Qy	230	PheThrPheGlySerGlyThrLysLeuGlnIleLysAla-----	242
Db	778	TGGACGTTCCGTTGAGGACCAACACTCGAGTTTCTAGAACAGAAATGATCTCCGAGAA	837
Qy	243	-----LeuGlnIleSerAsnSerValMetTyrPheSerSerValAlaProValLeu	259
Db	838	GATCTGAACCTAGAGATCAGCAACTCGGTGATGTACTTCAAGTTCTGTCTGCCAGTCTT	897
Qy	260	GlnLysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisPro	279
Db	898	CAGAAAGTGAATCTTACTTACTACCAACACGAGTGTGGATCTCCCTCACTGTGACACTT	957
Qy	280	ThrGlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGly	299
Db	958	ACCGGACATCTCAGCCCCAGAGACAGAAAGTTGTGGCCCCCTGGCTCAGTGAAGGG	1017
Qy	300	ThrGlyLeuAspPheLeuGlnAspProLysLeuCysTyrLeuLeuAspGlyIleLeuPhe	319
Db	1018	ACCGGATTTGAGCTTGATAGAGATCCCAACTGTGTAAGTGGAAATCCTCTTC	1077
Qy	320	IleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGlu	339
Db	1078	ATCTACGAGATCATCATCAACAGCCCTTACTGAGACAAATTCACAGAGATGACAG	1137
Qy	340	ThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrArgGlnLeuAsnLeuGlyArgArg	359
Db	1138	ACTCTCTGCAACTGCAAGACCCCAACAGCTTCAATAGTACTCAATCTTAGGGCGAAGA	1197
Qy	360	GlnGlnTyrAspValLeuGlnLysLysArgAlaArgAspProGlnMetGlyLysGln	379
Db	1198	GAGGAATATGACGCTTGGAGAAAGACGGGCTCGGATCCAGAGATGGAGGCAACAG	1257
Qy	380	GlnArgArgArgAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAla	399
Db	1258	CAGAGGAGGAGGAAACCCCAAGAGGAGTATACATCACTGACGCAAGAAAGACAAATGCGCA	1317
Qy	400	GlnAlaLysSerGlnIleGlyThrLysGlnGlnArgArgArgGlyLysGlnHisAspGly	419
Db	1318	GAAACCTTACAGTGAATCGGCACAAAGGCGAGAGGCGGAGAGGCAAGGGCAAGATGGC	1377
Qy	420	LeuTyrGlnGlnLysSerThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThr	439
Db	1378	CTTTACCAAGGCTCAGACATCGCAACCAAGACACCTATATGATGCCCTGCAATGCAAGCC	1437
Qy	440	LeuAlaProArg	443
Db	1438	CTGGCCCTCGC	1449





Db	1024	GGCGAAGAGAGGAAATATGACCTTTGGAGAGAGCCGGCTCGGGATTCACAGATGGGA	1083
Qy	377	GlyysgnginaTargArgAsnProInglugIyValTYrAsnaIaleuGlnLYsAsp	396
Db	1084	GGCAAAACGACAGAGAGAGGAACCCACAGAAAGCATATACATGCACTGCAGAAAGAC	1143
Qy	397	LysMetAlaGluAlaTYrSerGluIleGlyThrLYsGlyGluArgArgGlyLYsGly	416
Db	1144	AACATGGCAGAGAGCCTTACAGTAGATCCGGCACMAAAGCCAGAGCCGAGAGGCAAGGGG	1203
Qy	417	HisAspGlyLeuTYrGlnGlyLeuSerThrIaThrLYsAspThrTYrAspAlaLeuHis	436
Db	1204	CACAGTGCCTTTACACAGGGTCTCAGCACTGCACACAGAGACCTATGATGCTCGAT	1263
Qy	437	MetGlnThrLeuAlaProArg	443
Db	1264	ATGCAGACCTCGGCCCCCTCGC	1284
RESULT 4			
AAAT77137	AAAT77137 standard; DNA; 3058 BP.		
XX	AAAT77137;		
AC	AAAT77137;		
XX	25-MAR-2003 (revised)		
DT	04-MAR-1998 (first entry)		
DT	04-MAR-1998 (first entry)		
XX	Single chain antigen hybrid receptor DNA.		
DE	Single chain antigen hybrid receptor DNA.		
XX	Hybrid receptor; single chain antigen; gene therapy; diagnosis;		
KM	signal conduction; receptor; control region; ss.		
XX	Synthetic.		
OS	Homo sapiens.		
XX	Homo sapiens.		
FH	Key	Location/Qualifiers	
FT	misc_signal	/tag= a	562..570
FT	misc_signal	/note= "Noted in specification"	
FT	CDS	/tag= b	723..2136
FT	CDS	/note= "This sequence contains an intron"	
FT	sig_peptide	/tag= d	723..861
FT	sig_peptide	/label= leader_sequence	
FT	intron	/note= "This sequence contains an intron"	
FT	intron	/tag= c	768..849
FT	mat_peptide	/tag= c	862..2133
FT	mat_peptide	/product= "single_chain_antigen_hybrid_receptor"	
FT	polyA_site	/tag= e	2514..2519
FT	polyA_site	/tag= f	
XX	W09720938-A2.		
XX	W09720938-A2.		
XX	12-JUN-1997.		
XX	03-DEC-1996;	96WO-DE002334.	
XX	03-DEC-1996;	96WO-DE002334.	
XX	05-DEC-1995;	95DE-01045351.	
XX	05-DEC-1995;	95DE-01045351.	
PA.	(UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.		
XX	(UYFR-) UNIV FREIBURG KLINIKUM ALBERT-LUDWIGS.		
F1	Mertelmann R, Kulmburg P, Rosenthal F;		
XX	Mertelmann R, Kulmburg P, Rosenthal F;		
XX	WPI; 1997-319784/29.		
DR	P-PSDB; AAW24025.		
XX	P-PSDB; AAW24025.		
FT	Cells with hybrid receptor having extracellular and intracellular regions		
FT	of different origins - useful in gene therapy and diagnosis of tumours.		

Example 3, Fig 4; 46pp; German.

This DNA sequence encodes a novel single chain antigen hybrid receptor (HR) and contains an extracellular domain specific for the hapten 4-hydroxy-5-iodo-3-nitrophenyl acetate (NIP) coupled to the CD8- $\alpha$  region and the transmembrane and signal-conducting intracellular parts of the CD3-zeta molecule. Such hybrid receptors comprises a receptor part localised on the outside of the cell and specific to a particular signal molecule and a receptor part originating from another receptor, localised on the inside of the cell and capable of setting off a signal inside the cell. The cell should also contain at least one other gene construct with a control region which can interact with the signal sent out by the hybrid receptor and thereby control expression of a transgene bound to this control region. Such cells are useful in gene therapy or for diagnostic purposes. (Updated on 25-MAR-2003 to correct PR field.)

Sequence 3058 BP; 689 A; 831 C; 853 G; 685 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	1.08e-109	Length:	3058
Score:	1501.50	Matches:	297
Percent Similarity:	76.29%	Conservative:	44
Best Local Similarity:	66.44%	Mismatches:	79
Query Match:	64.03%	Indels:	27
DB:	2	Gaps:	6

US-09-596-774-7 (1-443) x NAT77137 (1-3058)

QY 1 GlnIleGlnLeuValGlnSerGlyProGlyIleuLeuLysPheProGlyGluThrValIle 20  
DB 862 CAGGTCCAACTGCACGACGCTCTGGGGCTGAGCTTGTGAAGCCTGGGCTTCAGTAACTG 921  
QY 21 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIleGlnAla 40  
DB 922 TCCGCAAGGCTTGTGCTACACCTTCACCAAGCTACGTAGATGACCTGGTGAAGACAGG 981  
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTrpIleAsnThrSerThrGlyLysSerThrPhe 60  
DB 982 CCTGGACGAGGCCCTTGATGGATGGAAAGGATGGATCTTAATGATGGTGGTCACTAAGTAC 1041  
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerIleuThrSerAlaAsnThrAlaTyr 80  
DB 1042 AATGAAAGTTCAAGACCAAGGCCACACTGATCTGTAAACAAACCTTCAGACAGCTTAC 1101  
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTrpGlu 100  
DB 1102 ATGACGCTCAGCAGCCCTGACATCTGAGAGACTCTGCGCTATTATTGTGCAAGATACAT 1161  
QY 101 ValTyr--HisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSer 119  
DB 1162 TACTACGGTATGACCTTGTGACTACTCTGGGGCCAAAGGACCAACGCTCACCGTCTCTCA 1221  
QY 120 GlyGlyValGlySerGlyGlyGlyGlySerGlyGlyGlyLysSerAspIleGlnLeuThr 139  
DB 1222 GGTGAGAGCGGTTCAGCGAGGGGCTCTGGCGCTGGCGGATCTCAGGCTTTGGGACA 1281  
QY 140 GlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAla 159  
DB 1282 CAG--GAATCTGCATCTACCACATCACTACCTGGTGAACATGACACATCACTTGTGCTCA 1338  
QY 160 SerGlnAspValTyr-----AsnAlaValAlaTrpTyrGlnGlnLysProGlyGln 176  
DB 1339 AGTACTGGGGCTGTTCACACTAGTAAGTAATTAAGCCAACTGGGTCCAAAGAAAAACGATCAT 1398  
QY 177 SerProLysLeuLeuIleTyrSerAlaLysSerArgTyrThrGlyValProSerArgPhe 196  
DB 1399 TTATTTCATGGTCTAAATAGGTGTGTACCAACCAACCGAGCTCCAGATGTTCTCGCAGATTC 1458  
QY 197 ThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 216  
DB 1459 TCAGGCTTCCTGATGTGAGACAAAGGCTGCCCTCAACATCAAGGGGACAGACGTAGAGAT 1518  
QY 217 LeuAlaValTyrPheCysGlnGlnHisAspAsnArgThrProPheThrPheGlySerGlyThr 236

```

Db      1519 GAGCGAATATATTCTGTCTATGTATGACAGCAACCATGGGTGTGTGGTGGAGGAAACC 1578
Qy      237 LysLeuGluIleIysAlaLeuGluIleSerAsnSerValMetYrPheSerSerValVal 256
Db      1579 AAACGTACTGTC-----CTCGAG----- 1596
Qy      257 ProValIleuGlnIysValAsnSerThrThrIlysProValIleuArgThrProSerPro 276
Db      1597 -----GAATTGACTACTACTACCAAGCGAGTCTCGAATCTCTCACT 1638
Qy      277 ValHisProThrGlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySer 296
Db      1639 GTGCACCTTACCTCGGAGCATCTCAGCCCGAGACAGCAAGATTGTGGCCCGCTGCTCA 1698
Qy      297 ValIysGlyThrGlyLeuAspPheLeuGluAspProIysLeuCysTyrLeuLeuAspGly 316
Db      1699 GTGAAGGGAGCGGAGTGGACTTC-----GACCCCAAACTCTGCTACTCTGTAGATGGA 1752
Qy      317 IleuPheIleIyrglyValIleIleThrAlaLeuYrLeuArgAlaIysPheSerArg 336
Db      1753 ATCTCTTTCATCTACGGAGTATCATCATCAGCCCTGTACTGTAGACGAAATTCAGCAGG 1812
Qy      337 SerIleGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuYrAsnGluLeuAsnIleu 356
Db      1813 AGTCAGAGACTGCTGCCACCTCGAGGACCCCAACGAGCTCTCAATGAGCTCAATCTA 1872
Qy      357 GLyArgArgGluGluYrAspValIleuGlnIysIysAlaArgAspProGluMetGly 376
Db      1873 GGGGGAAGAGAGAAATATGACTGCTTGGAGAAAGCGGGCTCGGATCCAGAAATGAGAG 1932
Qy      377 GLyIysGlnIleArgArgAsnProGlnIleGlyValYrAsnAlaIleuGlnIysAsp 396
Db      1933 GGCAAAACGACGAGAGAGAGAGACCCCGAGAGAGCATATGCATGACCTGCAAGAAAGAC 1992
Qy      397 LysMetAlaGluAlaYrSerGluIleGlyThrIysGlyIleAlaGlyArgGlyIysGly 416
Db      1993 AACATGGCAGAAAGCTACAGTACAGATCGGCAAAAGCGAGCGAGCGAGCAAGGGG 2052
Qy      417 HisAspGlyLeuYrGlnGlyLeuSerThrAlaThrIysAspThrYrAspAlaLeuHis 436
Db      2053 CACGATGGCTTTTACAGGGCTCAGCATGCTGCCACCAAGAGACCTTATGATGCTTCAT 2112
Qy      437 MetGlnThrLeuAlaProArg 443
Db      2113 ATCGAGACCTCGGCCCTCGC 2133

RESULT 5
AAT42036
ID AAT42036 standard; DNA; 2070 BP.
XX
AC AAT42036;
XX
DT 29-JAN-1997 (first entry)
XX
DE Plasmid pSW202-5 fragment encoding scFv(FRP5)-ETA.
XX
KW Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KM epidermal growth factor; receptor; plasmid pSW202-5; cancer; therapy;
XX
KW antitumour; exotoxin A; ETA; ss.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key
FT CDS 1..1956 Location/Qualifiers
FT FT /*tag= a
FT sig_peptide 1..63
FT /*tag= b
FT /*note= "Ompa signal peptide"
FT mat_peptide 64..1953

```

```

FT      EP739984-A1.
XX      30-OCT-1996.
XX
XX      26-APR-1995; 95SEP-00106275.
XX
XX      26-APR-1995; 95SEP-00106275.
XX
PR      (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX      Wels W, Schmidt M, Groner B;
XX
XX      WPI; 1996-478748/48.
XX
XX      P-PSDB; AAM05156.
XX
PT      Bivalent fusion proteins that bind epidermal growth factor receptor or
PT      analogues - and comprise at least two different cell surface binding
PT      domain(s), useful for tumour therapy.
XX
XX      Example 7, Page 21-22; 52pp; English.
XX
XX      A fragment (AAT42036) of plasmid pSW202-5 codes for scFv(FRP5)-ETA
XX      (AAM05136), comprising the single-chain binding region of murine
XX      monoclonal antibody FRP5, which is specific for human epidermal growth
XX      factor receptor erbB-2, joined to exotoxin A (ETA). A fragment of plasmid
XX      pMW152-5 (see also AAT42034) encoding scFv(FRP5) (AAM05134) was ligated
XX      to pSW202-5, a plasmid contg. the pseudomonas aeruginosa PAK ETA gene, to
XX      produce pSW202-5. This can be utilised in the construction of bacterial
XX      expression vectors (see also AAT42037-44) encoding bivalent fusion
XX      proteins (AAM05137-44) useful as antitumour agents

SQ      Sequence 2070 BP; 402 A; 675 C; 628 G; 365 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7,58e-95 Length: 2070
Score: 1312.50 Matches: 284
Percent Similarity: 68.59% Conservativity: 13
Best Local Similarity: 65.59% Mismatches: 55
Query Match: 55.97% Indels: 83
DB: 2 Gaps: 8

US-09-596-774-7 (1-443) x AAT42036 (1-2070)
Qy      1 GlnIleGlnLeuValGlnSerGlyProGluLeuIysIysProGlyGluThrValIysIle 20
Db      115 CAGGTACAACTGACAGCATGCTGACCTGAACGTGAAGAACCTGGAGAGACAGTCAAGATC 174
Qy      21 SerCysIysAlaSerGlyTyrProPheThrAsnYrGlyMetAsnTrpValIysGlnAla 40
Db      175 TCCTGCAAGGCTCTGGGTRTCTTTCAACAACTATGAAATGAACTAGGTGAAGCAGGCT 234
Qy      41 ProGlyGlnGlyLeuIysTrpMetGlyTyrIleAsnThrSerThrGlyIleSerThrPhe 60
Db      235 CCAGGACAGGGTTTAAAGTGAATGGGCTGATTAACACCTCCACGTGAGAGCAACATTT 294
Qy      61 AlaAspAspPheIysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaYr 80
Db      295 GCTGATGACTTCAGAGGACGGGTTTGACTTCTTTGAAACCTGTGCCAACACTGCTAT 354
Qy      81 LeuGlnIleAsnAsnLeuIysSerGluAspMetAlaThrYrPheCysAlaArgTrpGlu 100
Db      355 TTGCAGATCAACAACCTCAAAAGTGAAGCAAGCTACATATTTCTGTGCCAAGATGGAG 414
Qy      101 ValYrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
Db      415 GTTTACCAAGGCTACAGTTCTTACCTGAGGCGCAAGGAGCAACGCGTACCGTTCTCTGGC 474
Qy      121 GLyIysIysSerGlyGlyIysSerGlyIysGlyIysSerAspIleGlnLeuThrGln 140
Db      475 GGTGGGGTTCTGTGGCGGTGGCTCCGGGGGTGGCGGTTCTGACATCCAGCTGACCCAG 534

```

```

QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLeuAspSer 160
DB 535 TCTCAAAATTCCTGCTCACTTCACTAGAGACAGAGGCTCAGACTACCTGCAAGCCAGT 594
QY 161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB 595 CAGAGTGTATATAGCTGTGCTGCTGATCAACAGAAACAGACCAATCTCTAACTT 654
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerIly 200
DB 655 CTGATTTACTGCGATCTCCCGGTACACTGAGTCCCTTCCTGCTTCACTGCGAGTGC 714
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaAspLeuAlaValTyr 220
DB 715 TCTGGGCGGATTTCACTTTCACCTACCTACAGCTGTCAGAGCTGAAAGACTGGCACTTAT 774
QY 221 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
DB 775 TTTCTGTAGCAACATTTTGTATCTCATTCAGTTCCGCTCGGGGCAAAATTGGAGATC 834
QY 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 260
DB 835 AAAGCTCTAGAGCA----- 848
QY 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
DB 849 -----CATTCACTACCACTACCTGTA 869
QY 281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgGlySerValLysGlyThr 300
DB 870 GGG---CGGACAGCTGGCGCGCTGAC-----CGCGACACAGGCTGCGACCTGCCGCT 920
QY 301 GlyLeuAspPheLeuGlnAspProLysLeuGlySerTyrLeuLeuAspGlyTyrLeuPheIle 320
DB 921 GGAACACTTTCACCCGCTCATCGCCAGCCGCGGCTGGGACAACTGAGCAGTGGGCTA 980
QY 321 TyrGlyValIleIleThrAlaLeuTyrTyrLeuArgAlaLysPheSer----- 335
DB 981 TCCGCTGACAGCGGCTGCTGCTCTTCACTGAGCGCGGCTGCTCATGTGAAACCAAGTCCA 1040
QY 336 -----ArgSerAlaGlnThrAla-----AlaAs 343
DB 1041 CCAGGTATCCGCAACCGCCCTGGCCAGCCCGGACGCGGCGGACCTGGGCGAAAGCAT 1100
QY 343 nLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgGlyGlnTyrAs 363
DB 1101 CCGCGAGCAGCCGAGCAGCGCCGCTGCTGCTGACCTT-GGCGCGCCGCGAGAG----- 1154
QY 363 pValLeuGlnLysLysArgAlaArgAspProGlnMetGlyGlyLysGlnGlnArgArgAr 383
DB 1155 -----CGAGCGCTTCTGCTCG-----CGAGGCGACCGGCAACGACGA-- 1190
QY 383 GAsnProGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSe 403
DB 1190 ----- 1190
QY 403 rGlnIleGlyThrLysGlyGlnArgArgGlyLys 415
DB 1191 -----GGCCGCGCGGCAACCGCCGAGTGTGAG 1220

```

## RESULT 6

ID AAT42044 standard; DNA; 2730 BP.

AC AAT42044;

DT 29-JAN-1997 (first entry)

XX Plasmid pMS246-5-5 encoding scFv2 (FRP5/FRP5)-ETA (version 3).

XX Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;

KM epidermal growth factor; receptor; plasmid pMS246-5-5; cancer;

```

KM exotoxin A; ETA; antitumour; ss.
XX Mus; gp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX Key Location/Qualifiers
FH CDS 1..2700
FT sig_peptide 1..63
FT /tag= a
FT /note= "Ompa signal peptide"
FT mat_peptide 64..2697
FT /tag= c
FT /product= "scFv2 (FRP5/FRP5)-ETA"
XX EP73984-A1.
XX 30-OCT-1996.
XX 26-APR-1995; 95EP-00106275.
XX 26-APR-1995; 95EP-00106275.
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX Wels W, Schmidt M, Groner B;
XX WP1; 1996-478748/48.
XX DR P-PSDB; AAM05144.
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
PT analogues - and comprise at least two different cell surface binding
PT domain(s), useful for tumour therapy.
XX Example 12; Page 43-45; 52pp; English.
XX A fragment (AAT42044) of bacterial expression plasmid pMS246-5-5 codes
CC for scFv2 (FRP5/225)-ETA (AAM05144), and is obtd. by ligating 2 copies of
CC DNA coding for the single-chain binding region of murine monoclonal
CC antibody FRP5 (see also AAT42034) and portions of Pseudomonas aeruginosa
CC exotoxin A. pMS246-5-5 can be utilised in the prodn. of bivalent fusion
CC protein in bacterial (esp. E. coli) host cells. Such fusion proteins (see
CC also AAM05138-43) bind the epidermal growth factor receptor and are
CC useful as antitumour agents
XX SQ Sequence 2730 BP; 567 A; 832 C; 798 G; 533 T; 0 U; 0 Other;
XX Alignment Scores:
Pred. No.: 1,06e-94 Length: 2730
Score: 1312.50 Matches: 284
Percent Similarity: 68.59% Conservative: 13
Best Local Similarity: 65.59% Mismatches: 55
Query Match: 55.97% Indels: 83
DB: 2 Gaps: 8
US-09-596-774-7 (1-443) x AAT42044 (1-2730)
QY 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysPheProGlyGlnThrValLysIle 20
DB 115 CAGGTACACAGCAGCAGTCTGACCTGAACTGAAAGAGCTGAGAGACAGTCAAGATC 174
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
DB 175 TCTGTCAAGGCTCTGGGTATCTTTCACAAACTATGAAATGAACTGGGTGAACAGAGCT 234
QY ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60
DB 235 CCAGGACAGGCTTAAAGTGAATGAGCTGATTAACCTTCACTGAGAGTCAACATTT 294
QY AlaAspAspPheLysGlyArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80

```

```

Db 295 GCTGATGACTTCAGAGGACGGTTTGACTTCTTTGGAAACCTTCGCAACACCTGCTAT 354
QY 81 LeuGlnIleAsnAsnLeuSerGlnuAspMetAlaThrTyPheCysAlaIleGlyTrpGlu 100
Db 355 TTGCAGATCAACACCTCAAAAGTGAAGACATGGCTTACTTATTTCTGTGCAAGATGGAG 414
QY 101 ValTyriSeglyTyraValProTyrTrpGlnGlnIleThrValThrValSerSerGly 120
Db 415 GTTTACCAAGGCTACGTTCTTACTGGGGCCAAAGGACACAGGTACACCGTTTCTCTGGC 474
QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerGlyGlySerGly 140
Db 475 GGTGGCGGTCTCGGTGGCGGTGGCTCCGCGGTGGCGGTCTGACATCAGCTGACCCAG 534
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLeuValSer 160
Db 535 TCTCAAAATTCCTCTCCACTTCACTAGAGACAGGGGTACAGATACCTGCAAGGCCAGT 594
QY 161 GlnAspValTyraAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 180
Db 595 CAGAGATGTAATATGATGCTGTGCTGTATCAACAGAAACAGAGCAATCTCTAAACTT 654
QY 181 LeuIleTyraSerAlaSerSerArgTyraGlyValProSerArgPheThrGlySerGly 200
Db 655 CTGATTACTCGGCACTCTCCCGGTACAGTGAAGTCCCTTCTCGCTTCACTGGCAGTGGC 714
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
Db 715 TCTGGCGCGGATTCACCTTCAACATCAGAGTGTGCGAGGCTGAAGACTTGGCAGTTAT 774
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyraLysLeuGlnIle 240
Db 775 TTCTGTCAAGCAACATTTTCGTACTTCATTCAGTTCGGGTGGGACAAATGTGAAGATC 834
QY 241 LysAlaLeuGlnIleSerAsnSerValMetTyraPheSerSerValProValLeuGln 260
Db 835 AAAGCTTAGACA----- 848
QY 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
Db 849 -----CCATCATCAACCATCACTCAAGA 869
QY 281 GlyThrSerGlnProGlnArgProGlnAspCysArgProArgLysSerValLysGlyThr 300
Db 870 GGG---CGGCAAGCCTGGCGCGCTGAC-----CGCGCACAGCGCCTGCCACCTGCGCT 920
QY 301 GlyLeuAspPheLeuGlnAspProLysLeuGlySerTyraLeuAspGlyIleLeuPheIle 320
Db 921 GGAAGACTTTCACCCGCTATCGCCAGCGCGCGCTGGGACACACTGAGCAGTGGGCTTA 980
QY 321 TyrglyValIle-IleThrAlaLeuTyraLeuArgAlaLysPheSer----- 335
Db 981 TCCGGTGAGCGGCGTGGTCCCTTACTGCGCGCGGAGTGTGTCAGGAACAGGTGCA 1040
QY 336 -----ArgSerAlaGluThrAla-----AlaAs 343
Db 1041 CCAGGTATCCGCAACGCTTGGCCAGCGCCGCGAGCGCGCGCACTGGGCGGAAGCGAT 1100
QY 343 nLeuGlnAspProAsnGlnLeuTyraAsnGlnLeuAsnLeuGlyArgArgGlnGluTyra 363
Db 1101 CCGGAGACGCGGAGCAGCGCCCTGCGCCCTTAACCT--GGCCGCGCGCGAGAG-- 1154
QY 363 pValLeuGlnLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGlnArgArg 383
Db 1155 -----CGAGGCTTCGTCCG-----GCAGGGACCGGACAGCAGCA-- 1190
QY 383 gAsnProGlnGlnValTyraAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSe 403
Db 1190 ----- 1190
QY 403 rGlnIleGlyThrLysGlyGluArgArgGlyLys 415
Db 1191 -----GGCCGCGCGCGCCAGCCCAAGCTGTGAG 1220

```

```

RESULT 7
AAQ28262
ID AAQ28262 standard; DNA, 2011 BP.
XX
AC AAQ28262;
XX
DT 24-OCT-2003 (revised)
DT 25-MAR-2003 (revised)
DT 11-FEB-1993 (first entry)
XX
DE Fv (FRP5) - ETA fusion gene.
XX
KW Monoclonal antibody; light chain; heavy chain; tumour; c-erbB-2;
KW variable region; ETA, ss.
XX
OS Pseudomonas aeruginosa; PAK.
XX
Key Location/Qualifiers
FH 1..63
FT sig_peptide
FT /*tag= a
FT /label= ompA_signal_peptide
FT 64..87
FT misc_RNA
FT /*tag= b
FT /note= "FLAG peptide and enterokinase cleavage site"
FT 97..453
FT misc_RNA
FT /*tag= c
FT /label= FRP5_heavy_chain_variable_domain
FT 454..498
FT misc_RNA
FT /*tag= a
FT /label= Linker
FT 499..822
FT misc_RNA
FT /*tag= a
FT /label= FRP5_light_chain_variable_domain
FT 826..1911
FT CBS
FT /*tag= a
FT /label= ETA_1574-1747
FT 1912..2012
FT 3'clip
FT /*tag= a
FT /label= ETA_3'_non-coding_region
XX
PN BP502812-A1.
XX
PD 09-SEP-1992.
XX
PF 27-JAN-1992; 92EP-00810056.
XX
PR 05-FEB-1991; 91EP-00810079.
XX
PA (CIBA ) CIBA GEIGY AG.
XX
PI Wels WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX WPI, 1992-302096/37.
XX DR P-PsDB; AAR26982.
XX
PT Recombinant antibodies directed to growth factor receptor C-ERBB-2 - for
PT diagnosing and treating tumours expressing C-ERBB-2 e.g. breast or
PT ovarian tumours.
XX
PS Disclosure; Page 47-52; 67pp; English.
XX
XX The sequences given in AAQ28261-3 encode part of the exotoxin A (ETA)
CC sequence corresponding to positions 1574-1747 of the full exotoxin A
CC sequence. These sequences were used in the construction of Fv(FRP5)-ETA
CC fusion genes. The ETA sequence was used as a marker gene so that E. coli
CC transformed with the fusion gene could be identified. The fusion genes
CC were expressed in E. coli and the antibody was extracted. This
CC recombinant antibody can be used for the qualitative and quantitative
CC determination of c-erbB-2. This can be used for monitoring or in-vivo
CC localisation of tumours overexpressing c-erbB-2. (Updated on 25-MAR-2003
CC to correct FN field.) (Updated on 24-OCT-2003 to standardise OS field)
XX

```

SQ Sequence 2011 BP; 383 A; 653 C; 620 G; 355 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.27e-94 Length: 2011  
 Score: 1309.50 Matches: 280  
 Percent Similarity: 68.36% Conservative: 16  
 Best Local Similarity: 64.67% Mismatches: 48  
 Query Match: 55.84% Indels: 91  
 DB: 2 Gaps: 8

US-09-596-774-7 (1-443) x AAQ28262 (1-2011)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValIle 20  
 DB 97 CAGGTACCACTGCAGCACTGAGCTGAACTGAAGAGCCTGAGAGACAGTCAAGTTC 156  
 QY 21 SerCysAlaSerGlyTyrProPheThrAntyrglyMetAsnTTPValIleGlnAla 40  
 DB 157 TCCTGCAAGGCGCTCGGGTATCTTTCACAAACTATGGAATGAACCTGGGTGAAGCAGGCT 216  
 QY 41 ProGlyGlnGlyLeuValSerTyrMetGlyTyrIleAsnThrSerThrGlyIleSerPhe 60  
 DB 217 CCAGGACAGGGTTTAAAGTGAAGGCTGATTAACACTCCACTGAGAGAGTCAACATTT 276  
 QY 61 AlaAspPhePheGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80  
 DB 277 GCTATGACTTCAAGGACGCTTTCATCTTCTTGGAACCTCTGCCAACACTGCCAT 336  
 QY 81 LeuGlnIleAsnLeuLeuSerGlyLeuAspMetAlaThrTyrPheCysAlaArgTyrGlu 100  
 DB 337 TTGCAGATGAACAACCTCAAAAGTGAAGACATGGCTACATATTTCTGCAAGATGGAG 396  
 QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValIleSerSerGly 120  
 DB 397 GTTATCAACAGGCTACGTTCTTACTGGGGCCAAAGGACACAGCTCACCGTTCTTGGC 456  
 QY 121 GlyIleGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140  
 DB 457 GGTGGCGGTTTGTGGGGTGGCTCCGGCGGTGGCTTGTGACATCACTGACACCCAG 516  
 QY 141 SerHisPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaAsnSer 160  
 DB 517 TCTCACAATTCCTGTCACCTTCACTAGAGACAGGCTCACATCACTGCAGAGCCAGT 576  
 QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnIlePyrGlnIleSerProIleu 180  
 DB 577 CAGGATGATGATTAATGCTGTGCTGATCAACAGAAACAGGACATCTCTAAACTT 636  
 QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
 DB 637 CTGATTTACTCGGACATCTCCGGGACACTGAGTCCCTTCTCGCTTCACTGCGAGTGC 696  
 QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220  
 DB 697 TCTGGGCGGATTTCACTTCAACATCAGAGTGGGAGGCTGAAGACTGGCAGTTTAT 756  
 QY 221 PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyThrIleLeuGluIle 240  
 DB 757 TTTCTGTCAGCAACATTTTCGTAATTCATTCAGTTCGGCTCGGGGACAAATTTGAGATC 816  
 QY 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValProValIleuGln 260  
 DB 817 AAAGCTCTAGAG----- 828  
 QY 261 LysValAsnSerThrThrThrLysProValIleuArgThrProSerProValHisProThr 280  
 DB 829 -----GGCGGAGAGCTGGCGCGCTGAC-GCGGACACAGGCTGCCACCTCCGCT 878  
 QY 281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValIleGlyThr 300  
 DB 879 GGAAGACTTTCACCGCTCATCCCA-----GCCGCGCGCTGGGAGACA 920  
 QY 301 GlyLeuAspPheLeuGluAspProLysLeuCysTyrIleLeuAspGlyIleLeuPheIle 320

DB 921 -----ACTGAGCAGTCCGCTA 938  
 QY 321 TyrGlyValIle-IleThrAlaLeuTyrIleuArgAlaLysPheSer----- 335  
 DB 939 TCCGTTGACAGGCGTGGCTGCTTCTTACCTGGCGCGGACTGTATGAAACAGGTGCA 998  
 QY 336 -----ArgSerAlaGluThrAla-----AlaAs 343  
 DB 999 CCAGGTATCCGCAACGCTTGGCCAGCCCGGACGGCGGCGGACACTGGGCGAAGCAT 1058  
 QY 343 nLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyIleArgGlnGluTyrAs 363  
 DB 1059 CCGGAGACGCGGAGAGAGGCGCGCTGGCCCTGACCT-GGCGCGCGCGAGAG----- 1112  
 QY 363 pValIleuGlnLysArgAlaArgAspProGluMetGlyGlySerGlnGlnArgArgAr 383  
 DB 1113 -----CGAGGCTTCTGCTCG-----GCAGGGGACCGGCAACGACA- 1148  
 QY 383 GAspProGlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSe 403  
 DB 1148 ----- 1148  
 QY 403 rGluIleGlyThrLysGlyGluArgArgArgLys 415  
 DB 1149 -----GGCGGCGCGGCAACCGCGAGCTGTGAG 1178  
 RESULT 8  
 ID AAT29413 standard; DNA, 1862 BP.  
 AC AAT29413;  
 XX 16-OCT-2003 (revised)  
 DT 19-AUG-1996 (first entry)  
 XX  
 DE DGAL4-DT-scfv (FRP5) gene in pSM55-GD5.  
 XX  
 KW Nucleic acid transfer system; gene transfer; gene therapy;  
 KW cell targeting; multidomain protein; vector; cancer; GAL4;  
 KW diphtheria toxin; single chain antibody; scfv; FRP5; ss.  
 XX  
 OS Saccharomycetes; cerevisiae.  
 OS Corynebacterium; diphtheriae.  
 OS Mus sp.  
 OS Chimeric.  
 OS Synthetic.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 1..1908  
 FT /"tag" a  
 FT /product= "GAL4-DT-scfv (FRP5) fusion protein"  
 FN MO9613599-A1.  
 PD 09-MAY-1996.  
 XX  
 PD 31-OCT-1995; 95WO-EP004270.  
 PF 01-NOV-1994; 94EP-00810627.  
 PR  
 XX  
 PA (WELS/) WELS W.  
 XX  
 PI Wele W. Pominaya J;  
 XX  
 DR WPI; 1996-239505/24.  
 DR P-PSDB; AAR95057.  
 XX  
 PT Nucleic acid transfer system for gene therapy. e.g. against cancer -  
 PT includes toxin translocation domain to target nucleic acid to specific  
 PT cell.  
 XX  
 PS Example 20; Page 84-86; 106pp; English.

XX Plasmid pSW5-GDS contains a fusion gene (AAT29412) coding for a  
 CC multidomain fusion protein (AAR95057) a FLAG epitope, amino acids 2-147  
 CC of yeast GAL4, amino acids 196-384 of diphtheria toxin (DT) and a single  
 CC chain antibody, scFv(FRP5), based on a monoclonal antibody raised against  
 CC the HER2 antigen of human tumour cells, all separated by spacer peptides.  
 CC The GAL4 moiety acts as a DNA binding domain, DT as a translocation  
 CC domain, and scFv(FRP5) as a ligand domain. The multidomain protein, which  
 CC can be expressed in *E. coli*, is used with an effector nucleic acid in a  
 CC novel nucleic acid transfer system suitable for gene therapy. The  
 CC effector nucleic acid includes the gene to be delivered and a cognate  
 CC structure for the GAL4 DNA binding domain. (updated on 16-OCT-2003 to  
 CC standardise OS field)  
 XX  
 SQ Sequence 1862 BP; 522 A; 444 C; 446 G; 450 T; 0 U; 0 Other;  
 Alignment Scores:  
 Pred. No.: 1,52e-94 Length: 1862  
 Score: 1308.00 Matches: 245  
 Percent Similarity: 97.64% Conservative: 3  
 Best Local Similarity: 96.46% Mismatches: 6  
 Query Match: 55.78% Indels: 0  
 DB: 2 Gaps: 0  
 US-09-596-774-7 (1-443) x AAT29413 (1-1862)  
 QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 20  
 DB 1090 CAGGTACACACTGCAGCAGCTCTGACCTGACCTGAAAGAGCTGGAGAGACAGTCAAGATC 1149  
 QY 21 SerCysIysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40  
 DB 1150 TCCTGCAAGGCGCTCTGGGTATCTTTCAACAATATGGAATGAATGGGTCAAGCAGGCT 1209  
 QY 41 ProGlyGlnGlyLeuLysLysTyrMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60  
 DB 1210 CCAGAGACAGGGTTAAAGTGGATGGCTGATTAACCTCCACTGAGAGATCAACATTT 1269  
 QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80  
 DB 1270 GCTGATGACTTCAAGGAGCGCTTTGACTTCTTTGAAACCTTGCACACACTGCTAT 1329  
 QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100  
 DB 1330 TTGCAGATCAACAACCTCAAAAGTGAAGACATGGCTCATATTTCTGCAAGATGGAG 1389  
 QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnIleLysThrValThrValSerSerGly 120  
 DB 1390 GTTTACCAAGGCTACGTTCTTACTGGGAGCCAGAGACACAGGTCCAGTTCCTGAGC 1449  
 QY 121 GlyGlyLysSerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140  
 DB 1450 GGTGGCGGTTCTGGTGGCGGTGGCTCCGGGGGTGGCGGTCTTGACATCCAGTCAACCAG 1509  
 QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160  
 DB 1510 TCTCACAAAATTCCTGCTCACTTCAGTAGAGACAGGGTCAACCTCAAGGCGCAGT 1569  
 QY 161 GlnAspValTyrAsnAlaValAlaIleTrpGlnGlnLysProGlyGlnSerProLysLeu 180  
 DB 1570 CAGAGATGTGTAATGCTGTTGCTGGATTCACAGAAACCAAGACATCTCTAAACTT 1629  
 QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
 DB 1630 CTGATTTACTCGGATCTCCCGGTACACAGGATCCCTTCTGCTTCACTGGCAGTGGC 1689  
 QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220  
 DB 1690 TCTGGGCGGATTTCACTTTCACCATCAGCAGTGTGAGGCTCAAGACCTGCGCAGTTTAT 1749  
 QY 221 PheCysGlnGlnHisAspArgTyrProPheThrPheGlySerGlyThrLysLeuGluIle 240  
 DB 1750 TTCTGTACAGCACTTTTCTGATCTCCATTCAGTTCGGCTCGGGAGACAAATTTGAGATC 1809

QY 241 LysAlaLeuGlnLysSerAsnSerValMetTyrPheSerSer 254  
 DB 1810 AAAGCTTAAAGAGATCTCTGAGTGAAGAAAGATTTTCAGCC 1851  
 RESULT 9  
 AAT29412  
 ID AAT29412 standard; DNA; 1919 BP.  
 AC AAT29412;  
 XX  
 XX 16-OCT-2003 (revised)  
 DT 19-AUG-1996 (first entry)  
 XX  
 DE DGL4-DT-scfv(FRP5) gene in pSW50-GDS.  
 XX  
 XX Nucleic acid transfer system; gene transfer; gene therapy;  
 KW cell targeting; multidomain protein; vector; cancer; GAL4;  
 KW diphtheria toxin; single chain antibody; scFv; FRP5; ss.  
 XX  
 OS Escherichia; coli.  
 OS Saccharomyces; cerevisiae.  
 OS Corynebacterium; diphtheriae.  
 OS Mus sp.  
 OS Chimeric.  
 OS Synthetic.  
 OS  
 FH Key Location/Qualifiers  
 FT CDS 1..1908 /\*tag= a  
 FT sig\_peptide 1..63 /\*tag= b  
 FT mat\_peptide 64..1905 /\*tag= c  
 FT /\*tag= c /product= "GAL4-DT-scfv(FRP5) fusion protein"  
 FT  
 XX  
 XX W09613599-A1.  
 XX  
 PD 09-MAY-1996.  
 XX  
 PF 31-OCT-1995; 95WO-EP04270.  
 XX  
 XX 01-NOV-1994; 94EP-00810627.  
 PA (WEL6//) WEL5 W.  
 PI Wels W, Fominaya J;  
 XX  
 DR WPI; 1996-239505/24.  
 DR P-PSDB; AAR95056.  
 XX  
 PT Nucleic acid transfer system for gene therapy, e.g. against cancer -  
 PT includes toxin translocation domain to target nucleic acid to specific  
 PT cell.  
 XX  
 PS Example 19; Page 78-81; 106pp; English.  
 PS  
 CC Plasmid pSW50-GDS contains a fusion gene (AAT29412) coding for a  
 CC multidomain fusion protein (AAR95056) comprising an *E. coli* ompa signal  
 CC peptide, a FLAG epitope, amino acids 2-147 of yeast GAL4, amino acids 196  
 CC 384 of diphtheria toxin (DT) and a single chain antibody, scFv(FRP5),  
 CC based on a monoclonal antibody raised against the HER2 antigen of human  
 CC tumour cells, all separated by spacer peptides. The GAL4 moiety acts as a  
 CC DNA binding domain, DT as a translocation domain, and scFv(FRP5) as a  
 CC ligand domain. The multidomain protein, which can be expressed in *E.*  
 CC coli, is used with an effector nucleic acid in a novel nucleic acid  
 CC transfer system suitable for gene therapy. The effector nucleic acid  
 CC includes the gene to be delivered and a cognate structure for the GAL4  
 CC DNA binding domain. (updated on 16-OCT-2003 to standardise OS field)  
 XX  
 SQ Sequence 1919 BP; 534 A; 459 C; 462 G; 464 T; 0 U; 0 Other;  
 Alignment Scores:

Pred. No.: 1.58e-94 Length: 1919  
 Score: 1308.00 Matches: 245  
 Percent Similarity: 97.64% Conservative: 3  
 Best Local Similarity: 96.46% Mismatches: 6  
 Query Match: 55.78% Indels: 0  
 Gaps: 0

US-09-596-774-7 (1-443) x AAT29412 (1-1919)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValIle 20  
 DB 1147 CAGGTACACGACGACGACCTGACCTGAAAGAGCCTGGAGACAGTCAAGATC 1206  
 QY 21 SerCysValAsaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 40  
 DB 1207 TCCGTGCAAGGCTCTGGGTATCTTTCACAAACATGAAATGAACTGGGTGAAGCAGCT 1266  
 QY 41 ProGlyGlnGlyLeuValSerTyrMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60  
 DB 1267 CCAGGACAGGGGTTTAAAGTGGATGGGCTGATTAACACTCCACTGGAGAGTCAACATTT 1326  
 QY 61 AlaAspPhePheLeuValSerGlyPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80  
 DB 1327 GCTGATGACTTCAAGGAGCGGTTGACTTCTTGGAAACCTCTGCCAACACTGCCCTAT 1386  
 QY 81 LeuGlnIleAsnLeuValSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGlu 100  
 DB 1387 TTGCAGATCAACAACTTCAAAAGTGAAGACATGGCTCAATTTCTGTGCAAGATGGAG 1446  
 QY 101 ValTyrHisGlyTyrValProTyrTyrPylGlnGlyThrThrValThrValSerSerGly 120  
 DB 1447 GTTACACAGCGCTACGTTCTTACCTGGGGCCAAAGGACCAAGTCAAGCTTCTCTGGC 1506  
 QY 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140  
 DB 1507 GGTGGCGGTTCTGGTGGCGGTGGCTCCCGCGGTGGCTTGTGACATCCAGCTGACCCAG 1566  
 QY 141 SerHisValPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAsaSer 160  
 DB 1567 TCTCAACAATTCCTGCTCCACTTCACTAGAGAGACAGGAGTCACTGCAAGGCCAGT 1626  
 QY 161 GlnAspValTyrAsnAlaValAlaTyrPylGlnGlnValProGlyGlnSerProValLeu 180  
 DB 1627 CAGATGTGTAAATGCTGTTGCCGTATCAACAGAAACAGGACATCTCTTAAACTT 1686  
 QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
 DB 1687 CTGATTTACTCGCATCTCCCGGTACACTGAGTCCCTTCTCGCTTCACTGGCAGTGGC 1746  
 QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220  
 DB 1747 TCTGGGCGGATTCATCTTACCATCAGAGTGGCTGAAGAGCTGGAGCTGAGTTAT 1806  
 QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrThrLeuGluIle 240  
 DB 1807 TTCTGTGACGACATTTTTCGATCTCATTCACGTTCCGCTGGGGACAAATTTGAGATC 1866  
 QY 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSer 254  
 DB 1867 AAAGCTCTAGAGATCTCTCGAGTGAGAGAGATTTTCAGCC 1908

RESULT 10  
 AAT42038  
 ID AAT42038 standard; DNA; 2214 BP.

XX AAT42038;  
 DT 29-JAN-1997 (first entry)

XX Plasmid pMS238-5-TGF fragment encoding scFv(FRPS)/TGF alpha-ETA.  
 DE Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;  
 KW epidermal growth factor; receptor; plasmid pMS238-5-TGF; cancer;

KW exotoxin A; ETA; transforming growth factor alpha; TGF; antitumour; 88.  
 XX Homo; sapiens.  
 OS Mus; sp.  
 OS Pseudomonas; aeruginosa.  
 OS Synthetic.  
 OS Chimeric.  
 XX Key  
 FH CDS  
 FT Location/Qualifiers  
 FT 1..2100  
 FT /\*tag= a  
 FT sig\_peptide 1..63  
 FT /\*tag= b  
 FT /note= "OmpA signal peptide"  
 FT mat\_peptide 64..2097  
 FT /\*tag= c  
 FT /product= "scFv(FRPS)/TGF alpha-ETA"

XX EP739984-A1.  
 XX 30-OCT-1996.  
 XX 26-APR-1995; 95EP-00106275.  
 XX 26-APR-1995; 95EP-00106275.  
 XX (SANT-) SAN TUMORFORSCHUNGS GMBH.  
 XX Wels W, Schmidt M, Groner B;  
 DR WPI; 1996-478748/48.  
 DR P-PSDB; AAM05138.

PT Bivalent fusion proteins that bind epidermal growth factor receptor or  
 PT analogues - and comprise at least two different cell surface binding  
 PT domain(s), useful for tumour therapy.

XX Example 10; Page 25-27; 52pp; English.

XX A fragment (AAT42038) of bacterial expression plasmid pMS238-5-TGF codes  
 CC for scFv(FRPS)/TGF alpha-ETA (AAM05138), comprising the single-chain  
 CC binding region of murine monoclonal antibody FRPS (specific for human  
 CC epidermal growth factor receptor erbB-2, see also AAT42034) joined to  
 CC exotoxin A from Pseudomonas aeruginosa and human transforming growth  
 CC factor (TGF) alpha (see also AAM05137). pMS238-5-TGF can be utilised in  
 CC the prodn. of a bispecific fusion protein in bacterial (esp. E. coli)  
 CC host cells. Such fusion proteins (see also AAM05139-44) are useful as  
 CC antitumour agents

XX Sequence 2214 BP; 431 A; 710 C; 665 G; 408 T; 0 U; 0 Other;

XX Alignment Scores:

Pred. No.: 1.88e-94 Length: 2214  
 Score: 1308.00 Matches: 275  
 Percent Similarity: 71.19% Conservative: 24  
 Best Local Similarity: 65.48% Mismatches: 59  
 Query Match: 55.78% Indels: 64  
 Gaps: 5

US-09-596-774-7 (1-443) x AAT42038 (1-2214)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValIle 20  
 DB 115 CAGGTACACGACGACGACCTGACCTGAAAGAGCCTGGAGACAGTCAAGATC 174  
 QY 21 SerCysValAsaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 40  
 DB 175 TCCGTGCAAGGCTCTGGGTATCTTTCACAAACATGAAATGAACTGGGTGAAGCAGCT 234  
 QY 41 ProGlyGlnGlyLeuValSerTyrMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60  
 DB 235 CCAGGACAGGGGTTTAAAGTGGATGGGCTGATTAACACTCCACTGGAGAGTCAACATTT 294

QY	61	AlaAspAspPheIysGlyArgPheAspPheSerLeuGlnThrSerAlaSerThrAlaTyr	80
Db	295	GCTGATGACTTCACAGGACCGATTGACTTCTCTTGGAAACCTCTGCACACTGCCTAT	354
QY	81	LeuGlnIleAsnAsnLeuIysSerGluAspMetAlaThrTyrPheCysAlaArgTProGlu	100
Db	355	TTCGACATCAACAACTCAAAAGTGAAGACATGGCTACATATTCTGTGCAGATGGAG	414
QY	101	ValTyrHisGlyTyrValProTyrTTProGlyGlnGlyThrThrValThrValSerSerGly	120
Db	415	GTTTACCAGCGCTACGTTCTCTTACTGGGCAAGGAGACACGGTCAACGTTCTCTTGC	474
QY	121	GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln	140
Db	475	GGTGGCGGTTCGTGGCGGTGGCTCCGGCGGTGGCGGTTCACATTCACGTGACCGG	534
QY	141	SerHisIysPheLeuSerThrSerValGlyAspArgValSerIleThrCysIysAlaSer	160
Db	535	TCTCACTCAAAATTCCTGTCCACTTCAGTAGGAGACAGGGTACGATCACTCCGAAAGCCAGT	594
QY	161	GlnAspValTyrAsnAlaValAlaTyrProGlnGlnIleProGlyGlnSerProIysLeu	180
Db	595	CAGCATGTGTATATATCTGTTCCTGGTATTCACAGAAACAGAGAAATCTCTAACTT	654
QY	181	IleuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly	200
Db	655	CTGATTACTCCGCATCTCTCCGGTACACTGGAGTCTCTTCGCTTCACTGCAGTGGC	714
QY	201	SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr	220
Db	715	TCTGGCCGCGATTTCATCTTCACATCAAGAGTGTCAAGCTGTGAAGACTGTGACATTAT	774
QY	221	PheCysGlnGlnIleAspPheArgThrProPheThrPheGlySerGlyThrIysLeuGlnIle	240
Db	775	TTCTGTCAAGCAACATTTTGTGATCTCATTCACGTTCCGCTCGGGGACAAATTTGAGATC	834
QY	241	LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln	260
Db	835	AAAGCTCTAGAGCA-----	848
QY	261	LysValAsnSerThrThrThrIysProValLeuArgThrProSerProValHisProThr	280
Db	849	-----CAATCAATCAACATCACTTGA	869
QY	281	GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValIysGlyThr	300
Db	870	GGG---CGGACGCTGGCGCGCTGTAC-----CGGCAACAGCGCTGCACCTCCGCT	920
QY	301	GlyLeuAspPheLeuGlnAspProIysIleuCysTyrIleLeuAspGlyTyrIleLeuPheIle	320
Db	921	GGAGACTTTCACCCGCTCATTCGCAACCCGCGCGCTGGGAAACAATGTGAGCACTGGCGCTA	980
QY	321	TyrGlyValIleIleThrAlaLeuTyrIleuArgAlaIysPheSerAspSerAlaGluThr	340
Db	981	TCCGGTGCAGCGGCTGTGTCCCTTCACTGGCGGCGGAGCTGTCA-----	1022
QY	340	AlaIleAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyValArgGly	360
Db	1027	-----TGGAACCAAGGTCGACCA-GGT-----	1048
QY	360	GluTyrAspValLeuGlnIlyIysArgAlaArgAspProGluMetGlyGlyIysGlnGly	380
Db	1047	-----GATCCGCAAGCCCTGGCCAGCCCG	1072
QY	380	ArgArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnIlyAspIysMetAlaG	400
Db	1073	GCAGCGGCGGACCTTCGGCGCAAGGATCCGCGAGCAACCGGAGCAGCGCCGTCTGGGCC	1122
QY	400	IuAlaTyrSerGlnIleGlyThrIysGlyIuArgArgArgGlyIyIysIyHisAsp	418
Db	1133	TGACCCTTGCCGCGCGAGAGCGAGCCGCTTCGCGGAGGAGCAACCGGCAACGAC	1188

AAAT2039	
ID	AAAT2039 standard; DNA; 2793 BP.
XX	
AC	AAAT2039;
XX	
DT	29-JAN-1997 (first entry)
XX	
D2	Plasmid pMS238-5-225 fragment encoding scFv2(FRP5/225)-ETA.
XX	
KW	Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KM	epidermal growth factor; receptor; plasmid pMS238-5-225; cancer;
KX	exotoxin A; ETA; antitumour; ss.
OS	Mus; SP.
OS	Pseudomonas; aeruginosa.
OS	Synthetic.
OS	Chimeric.
FT	
FT	Key
FT	CDS
FT	Location/Qualifiers
FT	1..2679
FT	/tag= a
FT	sig_peptide
FT	1..63
FT	/tag= b
FT	/note= "Ompa signal peptide"
FT	mat_peptide
FT	64..2676
FT	/tag= c
FT	/product= "scFv2(FRP5/225)-ETA"
PN	
PN	EP739984-A1.
PD	
PD	30-OCT-1996.
XX	
PF	
PF	26-APR-1995; 95EP-00106275.
XX	
PR	
PR	26-APR-1995; 95EP-00106275.
XX	
PA	(SANT-) SAN TUMORFORSCHUNGS GMBH.
XX	
PI	
PI	Weis W, Schmidt M, Groner B;
DR	
DR	WPI; 1996-478748/48.
DR	P-PsDB; AA05139.
PT	
PT	Bivalent fusion proteins that bind epidermal growth factor receptor or analogues - and comprise at least two different cell surface binding domain(s), useful for tumour therapy.
XX	
PS	
PS	Example 11; Page 28-30; 52pp; English.
XX	
CC	A fragment (AAAT2039) of bacterial expression plasmid pMS238-5-225 codes
CC	for scFv2(FRP5/225)-ETA (AA05139), and is obtd. by ligating DNA
CC	fragments coding for the single-chain binding region of murine monoclonal
CC	antibody FRP5 (see also AAAT2034). portions of Pseudomonas aeruginosa
CC	exotoxin A and the single-chain binding region of murine monoclonal
CC	antibody 225 (see also AAAT2033). pMS238-5-225 can be utilised in the
CC	prodn. of bivalent fusion protein in bacterial (esp. E. coli) host cells.
CC	Such fusion proteins (see also AA05138-44) bind the epidermal growth
CC	factor receptor and are useful as antitumour agents
XX	
SO	Sequence 2793 BP; 593 A; 847 C; 806 G; 547 T; 0 U; 0 Other;
Alignment Scores:	
Pred. No.:	2..49e-94 Length: 2793
Score:	1308..00 Matches: 275
Percent Similarity:	71.19% Conservative: 24
Best Local Similarity:	65.48% Mismatches: 59
Query Match:	55.78% Indels: 64
DI:	Gaps: 5



Db	115	AAGGTACAACGTGCACAGCTGTGGACTGAACTGAAAGAGCCGTGGAGAGCACTCAAGACT	174
Qy	21	SerCybLysValAspSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla	40
Db	175	TCCTGCAGACGGCCTCTGGGTATCTCTTTCACAAACTATGAGATGAATGAGCGGTGAAGAGCGCT	234
Qy	41	ProGlyGlnGlyLeuLysTrpMetClyTrpIleAsnThrSerThyGlyGlnSerThrPhe	60
Db	235	CCAGAGACAGGGGTTTAAAGTGAATGGAGTGGACGTGAATTAACCTTCACCTGAGAGATCAACATT	294
Qy	61	AlaAspAspPheLysGlyYArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr	80
Db	295	GCTGATGACTTCACAGGACGGTTTATGACTCTCTTTGGAAACCTCTGCGCAACCTGCTCCT	354
Qy	81	LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTrpGlu	100
Db	355	TTGCAGATCAACCAACCTCAAAAGTGAAGACATGGCTACATATTTCTGTGCAGATGGGAG	414
Qy	101	ValTyrHisGlyTyrTrpValProTyrTrpGlyGlnGlyTyrThrValThyValSerGly	120
Db	415	GTTTACCAACGGCTACGGTCTTACTTACATGGGGCCAAAGGACACACGGTACCGTTCCTGTGC	474
Qy	121	GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln	140
Db	475	GGTGGCGGTTCTGGTGGCGGTGGCTCCGGCGGTGGCGGTTCTGCAATCACTGACGCCAG	534
Qy	141	SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer	160
Db	535	TCTCACAAAATTCCTGTCCACTTCAGTAGGAGACAGGTCAGCATCACTGCAGAGGCCAGT	594
Qy	161	GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu	180
Db	595	CAGATGTGTATTAAGCTGTGGCTGGATCAACAGAAACCGAGCAATCTCTAACTT	654
Qy	181	LeuIleTyrSerAlaSerSerSerArgTyrThrGlyValProSerArgPheThyGlySerGly	200
Db	655	CTGATTTTCTGGGACTCTCCCGGATGACTGAGATCCCTTCTGCTTCACTGGCACTGGC	714
Qy	201	SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr	220
Db	715	TCTGGGGCGGATTCATTCACTTCACATCAAGAGTGTGCAGGGCTGGAAGACTGCACATTAT	774
Qy	221	PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrLysLeuGluIle	240
Db	775	TTCTGTCAAGCAACATTTTCGTACTCACTTCACTTCAGTTCGGCTCGGGAGACAAATTCGAGATC	834
Qy	241	LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln	260
Db	835	AAACCTCTAGAGCA-----	848
Qy	261	LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr	280
Db	849	-----CCATCACTACCATCACTCACTGA	869
Qy	281	GlyThrSerGlnProGlnArgProGlyLysAspCysArgProArgGlySerValLysGlyThr	300
Db	870	GGG---CGGCAAGCTCGCCGCGCTGAC-----CGCGACACAGGCGCTGCACCTCCGCT	920
Qy	301	GlyLeuAspPheLeuGluAspProLysLysCysTyrIleLeuAspGlyTyrIleLeuPheIle	320
Db	921	GGAAACTTTCACCCGTCATCGCCACGACCGCGCGCTGGAGAACATCGAGACAGTGGCGCTA	980
Qy	321	TyrGlyValIle-IleThrAlaLeuTyrIleuArgAlaLysPheSerArgSerAlaGluThr	340
Db	981	TCCGGTGAAGGGCGTGTGCGCTTACTCTGCGCGCGACCTGTA-----	1022
Qy	340	AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyYArgArgGly	360
Db	1027	-----TGAAACAGGTGCACCA-GGT-----	1046
Qy	360	uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGluMetGlyGlyLysGlnGly	380
Db	1047	-----GATCCGACAGCGCCCTGGCCACCCCG	1077

Gy	380	naryArgrtgrgAnP-ro-GingluGlyValAlTYrAsnAlaleuGlnLYsAapLYeMeTalaG	400
Dd	1073	GCAAGCGGGCAGACTTGCGCCGAAGCGANTCCGCAGACGCCGGAAGCAAGCCCCGTCTGCCC	1133
Gy	400	IuaLaTySeScIuileGlyThrLySglYGuArGrArGlYlYsgLYHIsASP	418
Dd	1133	TGACCCTTGCCGCGCCGAGAGCGAGCGCTTCGTCGGGAGGGGACCCGGCAACGAC	1188
<b>RESULT 12</b>			
ID	AAT42043	standard; DNA; 2793 BP.	
AC	AAT42043;		
DT	29-JAN-1997	(first entry)	
DE	Plasmid pMS238-5-5 encoding scFv2(FRP5/FRP5)-ETA (version 2).		
KW	Single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;		
KV	epidermal growth factor; receptor; plasmid pMS238-5-5; cancer;		
KM	exotoxin A; ETA; antitumour; se.		
XX	Mus; sp.		
OS	Pseudomonas; aeruginosa.		
OS	Synthetic.		
OS	Chimeric.		
FH	Key	Location/Qualifiers	
FT	CDS	1..2679	
FT		/*tag= a	
FT	sig_peptide	1..63	
FT		/*tag= b	
FT		/note= "OmPa signal peptide"	
FT	mat_peptide	64..2676	
FT		/*tag= c	
FT		/product= "scFv2(FRP5/FRP5)-ETA"	
PN	EP739984-A1.		
PD	30-OCT-1996.		
XX			
PF	26-APR-1995;	95EP-00106275.	
PR	26-APR-1995;	95EP-00106275.	
PA	(SANT-) SAN TUMORFORSCHUNGS GMBH.		
PI	Weis W, Schmidt M, Groner B;		
DR	WP1: 1996-478748/48. P-PSDB; AAM05143.		
PT	Bivalent fusion proteins that bind epidermal growth factor receptor or analogues - and comprise at least two different cell surface binding domain(s), useful for tumour therapy.		
PS	Example 12; Page 40-42; 52pp; English.		
XX			
CC	A fragment (AAT42043) of bacterial expression plasmid pMS238-5-5 codes for scFv2(FRP5/225)-ETA (AAM05143), and is obtd. by ligating 2 copies of DNA coding for the single-chain binding region of murine monoclonal antibody FRP5 (see also AAT42034) and portions of Pseudomonas aeruginosa exotoxin A. pMS238-5-5 can be utilised in the prodn. of bivalent fusion protein in bacterial (esp. E. coli) host cells. Such fusion proteins (see also AAM05138-44) bind the epidermal growth factor receptor and are useful as antitumour agents		
SO	Sequence 2793 BP; 578 A; 858 C; 812 G; 545 T; 0 U; 0 Other;		
<b>Alignment Scores:</b>			
Pred. No.:	2.49e-94	Length:	2793
Score:	1308.00	Matches:	275

Alignment Scores:	
Pred. No.:	2,49e-94
Score:	1308.00
	Length: 2793
	Matches: 275

Percent Similarity: 71.19%  
 Best Local Similarity: 65.48%  
 Query Match: 55.78%  
 DB: 2  
 Conservative: 24  
 Mismatches: 59  
 Indels: 64  
 Gaps: 5

US-09-596-774-7 (1-443) x AAT42043 (1-2793)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGluValThrValIle 20
DB 115 CAGGATCAACAGCAGCAGTCTGACCTGAACTGAAAGAGCCTGGAGAGACGTCAAGATC 174
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 40
DB 175 TCCTGCAAGCGCTCGGGTATCTCTTCACAACCTAGAAAGAACTGGGTGAGAGAGCT 234
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrTrpIleAsnThrSerThrGlyLysThrPhe 60
DB 235 CCAGGACAGGGTTTAAAGTGAAGTGGCTGATTAACACCTCCACTGGAGAGTCAACATTT 294
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 295 GCTATGATGCTTCAAGGAGCGGTTTACTTCTTGGAAACCTCTGCCAACACTGCCCTAT 354
QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
DB 355 TTGCAGATCAACAACCTCAAAAGTGAAGACATGGCTACATATTTCTGTGCAAGATGGAG 414
QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerGly 120
DB 415 GTTACCAACGGCTACGTTCTTACTGGGGGACAGGAGACAGGTCACCGTTCTCTGGC 474
QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
DB 475 GGTGGCGGTTTGTGGCGGTTGCTCCGGCGGTGGCGGTTTGTGACATCCAGTACCAG 534
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer 160
DB 535 TCTCACAATTCCTGCTCACCCTTCACTGAGAGACAGGCTCAGCATCACCCTGCAAGCCAGT 594
QY 161 GlnAspValTyrAsnAlaValAlaTrpGlyGlnGlyLysProGlyGlnSerProLysLeu 180
DB 595 CAGAGATGTAATGCTGTTGCTGATCAACAGAAACAGACATCTCCCTAACTT 654
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 655 CTGATTACTCGGCATCTCCCGGTACACTGAGATCTCTTCTGCTTCACTGGCAGTGGC 714
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 715 TCTGGGCGGATTTCACTTTCACATCAGCAGTGTGAGGCTGAAAGACTGGCAGTTTAT 774
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGlnIle 240
DB 775 TTCTGTACGACAACTTTTCTGACTCCATTCACGTTCCGCTCGGGGACAAATTTGGAGATC 834
QY 241 LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
DB 835 AAAGCTCTAGAGCA----- 848
QY 261 LysValAsnSerThrThrLysProValLeuArgThrProSerProValHisProThr 280
DB 849 -----CCATCATCAGCATCAGCTAGA 869
QY 281 GlyThrSerGlnProGluAspProGluAspCysArgProArgLysValLysGlyThr 300
DB 870 GGG---CGGACGCTGGCGCGGTGAC-----CGCGACAGAGCCTGCCACCTGCCGCT 920
QY 301 GlyLeuAspPheLeuGluAspProLysLeuGlyTyrLeuLeuAspGlyIleLeuPheIle 320
DB 921 GGAGACTTTTCAACCGTCATCGCCAGCGCGGTGGGAAACACTGAGAGAGTGGCGCTA 980
QY 321 TyrGlyValIle-IleThrAlaLeuTyrLeuAlaGlyAlaLysPheSerArgSerAlaGlu 340
DB 340

```

```

DB 981 TCCGGTGACAGCGGTGGTCCGCTTACTGTCGGCGCCGCACTGTCA----- 1026
QY 340 rAlaAlaAsnLeuGlnAspProAsnGlnIleuTyrAsnGlnLeuAsnLeuGlyArgArgI 360
DB 1027 -----TGGAAACAGGTGAGCA--GGT----- 1046
QY 360 uGluTyrAspValLeuGlnLysLysArgAlaArgAspProGluMetGlyLysGlnI 380
DB 1047 -----GATCCGCAACGCCCTGGCCAGCCCCG 1072
QY 380 nArgArgArgAsnPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400
DB 1073 CGAGCGCGCGGACCTTGGCGGAGCGATCCGCGAGAGCGCGAGAGCGCGCTGTGCC 1132
QY 400 lualTyrSerGlnIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp 418
DB 1133 TGACCTTGCGCGCGCGCGAGAGGAGCGCTTGTCTCGCGAGCGAGCGCAACGAC 1188

RESULT 13
AAT42041
ID AAT42041 standard; DNA; 3177 BP.
XX
AC AAT42041;
XX
DT 29-JAN-1997 (first entry)
XX
DE plasmid pMS240-5-225 fragment encoding scFv2 (FRP5/225)-ETA.
XX
KW single chain antibody; scFv; monoclonal antibody; Mab; EGF; erbB-2;
KW epidermal growth factor; receptor; plasmid pMS240-5-225; cancer;
KW exotoxin A; ETA; antitumour; ss.
XX
OS Mus; sp.
OS Pseudomonas; aeruginosa.
OS Synthetic.
OS Chimeric.
XX
FH Key Location/Qualifiers
FT CDS 1..3063
FT sig_peptide /+tag= a
FT /+tag= b
FT /note= "OmpA signal peptide"
FT mat_peptide 64..3061
FT /+tag= c
FT /product= "scFv2 (FRP5/225)-ETA"

XX EP739984-A1.
XX
XX 30-OCT-1996.
XX
XX
XX PF 26-APR-1995; 95EP-00106275.
XX
XX PR 26-APR-1995; 95EP-00106275.
XX
XX (SANT-) SAN TUMORFORSCHUNGS GMBH.
XX
XX Wels W, Schmidt M, Groner B;
XX
XX WPI; 1996-4/78/48/48.
XX
XX DR P-PSDB; AAM05141.
XX
XX Bivalent fusion proteins that bind epidermal growth factor receptor or
XX PT analogues - and comprise at least two different cell surface binding
XX PT domain(s), useful for tumour therapy.
XX
XX Example 11; Page 34-36; 52pp; English.
XX
XX A fragment (AAT42041) of bacterial expression plasmid pMS240-5-225 codes
XX for scFv2 (FRP5/225)-ETA (AAM05141), and is obtd. by ligating DNA
XX fragments coding for the single-chain binding region of murine monoclonal
XX antibody FRP5 (see also AAT42034), portions of Pseudomonas aeruginosa
XX exotoxin A and the single-chain binding region of murine monoclonal

```

CC antibody 225 (see also AAT42033). pMS240-5-225 can be utilized in the  
 CC prodn. of bivalent fusion protein in bacterial (esp. E. coli) host cells.  
 CC Such fusion proteins (see also AAM05138-44) bind the epidermal growth  
 CC factor receptor and are useful as antitumour agents

XX Sequence 3177 BP, 654 A, 993 C, 939 G, 591 T, 0 U, 0 Other;

## Alignment Scores:

Pred. No.:	2,91e-94	Length:	3177
Score:	1308.00	Matches:	275
Percent Similarity:	71.19%	Conservative:	24
Best Local Similarity:	65.48%	Mismatches:	59
Query Match:	55.78%	Indels:	64
DB:	2	Gaps:	5

US-09-596-774-7 (1-443) x AAT42041 (1-3177)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValIle 20
   |||||
DB 115 CAGGTACAACTGCAGACAGTCTGACCTGAAGAAAGCTGGAGAGACAGTCAAGATC 174
QY 21 SerGlyValSerGlyTyrProPheThrSerGlyMetLeuThrValIleGln 40
   |||||
DB 175 TCCGTCAAGGCTCTGGGTATCTTTCACAACTATGAACTGGGTGAAAGCAGGCT 234
QY 41 ProGlyGlnGlyLeuValTyrMetGlyTyrIleThrSerThrGlyGluSerPhe 60
   |||||
DB 235 CAGGACAGGGGTAAAGTGGATGGCTGAGTAAACCTCCTCAGTGAAGTCAACATT 294
QY 61 AlaAspAspPheGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAla 80
   |||||
DB 295 GCTGATGACTTCAAGGAGCGCTTGACTTCTCTTGGAAACCTGCGCAACAGCGCAT 354
QY 81 LeuGlnIleAsnAsnLeuValSerGlyAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
   |||||
DB 355 TTGGAGATCAACAACCTCAAAAGTGAAGACATGCTCATTTCTGTGCAAGATGGAG 414
QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
   |||||
DB 415 GTTTACACAGGCTACGTTCTTACTGGGCGCAAGGACACGCTCCCTCTGGGC 474
QY 121 GlyGlyGlySerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
   |||||
DB 475 GGTGGCGGTTCTGGTGGCGGTGCTCCGCGCGTGTGATCATCAGCTACCCAG 534
QY 141 SerHisIlePheLeuSerThrSerValGlyAspArgValSerIleThrCysGlyAlaSer 160
   |||||
DB 535 TCTCACAAATTCCTGCTCAGTTCAGTAGAGACAGGCTCAGCATCAGCTGCAGGCGCAGT 594
QY 161 GluAspValTyrAsnAlaValAlaTrpTyrGlnGlnIleValSerProGlyGlnSerProGlyLeu 180
   |||||
DB 595 CAGAGTGTATATATGCTGTTGCTGGTATCAACAAGAAACAGGACAAATCTCTAACTT 654
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
   |||||
DB 655 CTGATTTACTCGGATCTCTCCCGGTACACGTGAATCCCTTCTCCCTTACCTGCGAGTGGC 714
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
   |||||
DB 715 TCTGGGCGGATTTCACTTTCACATCAGCAGTGTGAGGCTGAAAGACTGGCAGTTTAT 774
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrThrValLeuGluIle 240
   |||||
DB 775 TTCTGTAGCAACATTTTCTGACTTCATTCAGTTCGCTGGGAGCAAAATTTGAAATC 834
QY 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValValProValLeuGln 260
   |||||
DB 835 AAGGCTTAAAGCA----- 848
QY 261 LysValAsnSerThrThrThrLysProValLeuArgThrProSerProValHisProThr 280
   |||||
DB 849 -----CCATCATCATCAGCTACCTAGA 869

```

```

QY 281 GlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValIleGlyThr 300
   |||||
DB 870 GGG---CGGACGCTTGCGCGGCTGAC-----CGGACACAGCGCTGCCACCTGGCGCT 920
QY 301 GlyLeuAspPheLeuGluAspProLeuLeuCysTyrIleLeuAspGlyIleLeuPheIle 320
   |||||
DB 921 GGAGACTTTCACCCGTCATCGCCAGCGCGGCTGGAAACACTGAGAGAGTGGGCTA 980
QY 321 TyrGlyValIle-IleThrAlaLeuTyrIleArgAlaIlePheSerArgSerAlaGluThr 340
   |||||
DB 981 TCCGTGACAGCGGCTGGTGGCTCTTACCTGGCGCGGAGCTGTCA----- 1026
QY 340 rAlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgArgGly 360
   |||||
DB 1027 -----TGAACAGGTGACCA-GGT----- 1046
QY 360 uGluTyrAspValLeuGluValSerValArgAlaArgAspProGluMetGlyIleGlnGln 380
   |||||
DB 1047 -----GATCCGCAACGCCCTGGCCAGCCCG 1072
QY 380 nArgArgArgAspPro-GlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaG 400
   |||||
DB 1073 GCAGCGCGCGGACTGTGGCGAAGCGATCCGCGACAGCCGAGACAGCCGCTGGGCC 1132
QY 400 LuAlaTyrSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyHisAsp 418
   |||||
DB 1133 TGACCTGGCGCGCGCGAGAGAGCGCTTCCGCGGACAGGACCGGCAACAG 1188

```

## RESULT 14

AAT29409 ID AAT29409 standard; DNA, 1692 BP.

XX AC AAT29409;

XX XX 16-OCT-2003 (revised)

DT DT 18-AUG-1996 (first entry)

XX DE scfV(PRPS)-DETA-DGAL4 fusion gene in pWP46-5.

KW Nucleic acid transfer system; gene transfer; gene therapy;

KW cell targeting; multidomain protein; vector; cancer; exotoxin A; DETA;

XX KW ompA; signal peptide; single chain antibody; scfV; Gal4; ss.

OS Escherichia; coli.

OS Mus; sp.

OS Pseudomonas; aeruginosa.

OS Saccharomyces cerevisiae.

OS OS Chimeric.

OS OS Synthetic.

XX XX key

XX XX CDS

XX XX 09-MAY-1996.

XX XX 31-OCT-1995;

XX XX 01-NOV-1994;

XX XX (WELLS) WELLS W.

XX XX Wells W, Fominaya J;

XX XX WPI; 1996-239505/24.

XX XX P-PSDB; AAR35053.

Location/Qualifiers

1..1656

/\*tag= a

1..63

/\*tag= b

64..1653

/\*tag= c

/product= "scfV(PRPS)-DETA-DGAL4"

XX Nucleic acid transfer system for gene therapy, e.g. against cancer - PT includes toxin translocation domain to target nucleic acid to specific cell.

PS Example 8; Page 56-59; 106pp; English.

CC PlasmidMWp6-5 contains a fusion gene (AAR29409) coding for a  
CC multidomain fusion protein (AAR50503) comprising an E. coli ompA signal  
CC peptide, a FLAG epitope, FRP5 scfv (derived from anti-HER2 antigen  
CC monoclonal antibody FRP5), amino acids 232-366 of *Pseudomonas aeruginosa*  
CC exotoxin A (ETA), amino acids 2-147 of yeast Gal4, and an endoplasmic  
CC reticulum retention sequence, all separated by spacer peptides. The scfv  
CC acts as a ligand domain, ETA as a translocation domain, and Gal4 as a DNA  
CC binding domain. The multidomain protein, which can be expressed in *E.*  
CC coli, is used with an effector nucleic acid in a novel nucleic acid  
CC transfer system suitable for gene therapy. The effector nucleic acid  
CC includes the gene to be delivered and a cognate structure for the Gal4  
CC DNA binding domain. (Updated on 16-Oct-2003 to standardise OS field)

Sequence 1692 BP; 431 A; 446 C; 450 G; 365 T; 0 U; 0 Other;

### Alignment Scores:

Pred. No.:	2,35e-94	Length:	1692
------------	----------	---------	------

Score:	1305.00	Matches:	278
--------	---------	----------	-----

Percent Similarity:	68.36%	Conservative:	18
Best Local Similarity:	64.30%	Mismatch:	43

Best Local Similarity:	64.20%	Mismatches:	43
Query Match:	55.65%	Indels:	95

DB:	2	GAPs:	10
Query match:	25.0%	Indels:	93

1000

US-09-596-774-7 (1-443) X AAT29409 (1-1692)

[illegible]

Db	1715	1716	1717	1718	1719	1720	1721	1722	1723	1724	1725	1726	1727	1728	1729	1730	1731	1732	1733	1734	1735	1736	1737	1738	1739	1740	1741	1742	1743	1744	1745	1746	1747	1748	1749	1750	1751	1752	1753	1754	1755	1756	1757	1758	1759	1760	1761	1762	1763	1764	1765	1766	1767	1768	1769	1770	1771	1772	1773	1774	1775	1776	1777	1778	1779	1780	1781	1782	1783	1784	1785	1786	1787	1788	1789	1790	1791	1792	1793	1794	1795	1796	1797	1798	1799	1800	1801	1802	1803	1804	1805	1806	1807	1808	1809	1810	1811	1812	1813	1814	1815	1816	1817	1818	1819	1820	1821	1822	1823	1824	1825	1826	1827	1828	1829	1830	1831	1832	1833	1834	1835	1836	1837	1838	1839	1840	1841	1842	1843	1844	1845	1846	1847	1848	1849	1850	1851	1852	1853	1854	1855	1856	1857	1858	1859	1860	1861	1862	1863	1864	1865	1866	1867	1868	1869	1870	1871	1872	1873	1874	1875	1876	1877	1878	1879	1880	1881	1882	1883	1884	1885	1886	1887	1888	1889	1890	1891	1892	1893	1894	1895	1896	1897	1898	1899	1900	1901	1902	1903	1904	1905	1906	1907	1908	1909	1910	1911	1912	1913	1914	1915	1916	1917	1918	1919	1920	1921	1922	1923	1924	1925	1926	1927	1928	1929	1930	1931	1932	1933	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020	2021	2022	2023	2024	2025	2026	2027	2028	2029	2030	2031	2032	2033	2034	2035	2036	2037	2038	2039	2040	2041	2042	2043	2044	2045	2046	2047	2048	2049	2050	2051	2052	2053	2054	2055	2056	2057	2058	2059	2060	2061	2062	2063	2064	2065	2066	2067	2068	2069	2070	2071	2072	2073	2074	2075	2076	2077	2078	2079	2080	2081	2082	2083	2084	2085	2086	2087	2088	2089	2090	2091	2092	2093	2094	2095	2096	2097	2098	2099	2100	2101	2102	2103	2104	2105	2106	2107	2108	2109	2110	2111	2112	2113	2114	2115	2116	2117	2118	2119	2120	2121	2122	2123	2124	2125	2126	2127	2128	2129	2130	2131	2132	2133	2134	2135	2136	2137	2138	2139	2140	2141	2142	2143	2144	2145	2146	2147	2148	2149	2150	2151	2152	2153	2154	2155	2156	2157	2158	2159	2160	2161	2162	2163	2164	2165	2166	2167
----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

```

FT      /*tag= d
FT      /label= linker
FT      misc_RNA      491..814
FT      /*tag= e
FT      /label= FRPS_light_chain_variable_domain
FT      mat_peptide      815..2155
FT      /*tag= f
FT      /label= phoA_coding_region
XX      EP502812-A1.
XX      09-SEP-1992.
XX      27-JAN-1992; 92EP-00610056.
XX      05-FEB-1991; 91EP-00810079.
XX      (CIBA ) CIBA GEIGY AG.
XX      Wels WS, Hynes NE, Harwerth I, Groner B, Hardman N, Zwickl M;
XX      WPI: 1992-302096/37.
XX      P-PSDB; AAR26980.
XX      Recombinant antibodies directed to growth factor receptor C-ERBB-2 - for
XX      diagnosing and treating tumours expressing C-ERBB-2 e.g. breast or
XX      ovarian tumours.
XX      PS      Disclosure; Page 34-40; 67pp; English.
XX      CC      The sequence given encodes the single chain recombinant antibody
XX      CC      designated Fv(FRPS)-phoA. The alkaline phosphatase gene (phoA) was used
XX      CC      as a marker gene so that E. coli transformed with the fusion gene could
XX      CC      be identified. The fusion gene was expressed in E. coli and the antibody
XX      CC      was extracted. This recombinant antibody can be used for the qualitative
XX      CC      and quantitative determination of c-erbB-2. This can be used for
XX      CC      monitoring or in-vivo localisation of tumours overexpressing c-erbB-2.
XX      CC      (Updated on 25-MAR-2003 to correct PN field.)
SQ      Sequence 2233 BP; 566 A; 573 C; 606 G; 488 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.:      4,74e-94      Length:      2233
Score:      1303.00      Matches:      245
Percent Similarity:      95.00%      Conservative:      2
Best Local Similarity:      94.23%      Mismatches:      1
Query Match:      55.57%      Indels:      12
DB:      2      Gaps:      1

US-09-596-774-7 (1-443) x AAQ28257 (1-2233)
QY      1  GlnlleglneuvAlGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
DB      89  CAGGTACCACTGCAGCGTCTGCACCTGAACCTGAAGAGCTGGAGAGACAGTCAAGATC 148
QY      21  SerGlyLysAlaSerGlyTyrProPheThrAentYrGlyMetAentProValLysGlnAla 40
DB      149  TCCGCAAGGCGCTCTGGGTATCTTTCACAAACTATGGAATGAACCTGSGTGAAGCAGGCT 208
QY      41  ProGlyGlnGlyLeuLysTyrPheGlyTyrPheLeuThrSerThrGlnGlyLysThrPhe 60
DB      209  CCAGAGACAGGCTTTAACTGATGGCTGATTAACACCTCCACCTGAGAGATCAACATT 268
QY      61  AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAentThrAlaTyr 80
DB      269  GCTGATGACTTCAAGGAGCGGTGACTTCTTTGGAAACCTCGCCAAACACTGCGCTAT 328
QY      81  LeuGlnIleAentLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
DB      329  TTGCAGATCAACAACCTCAAAAGTGAAGACATGGCTACATATTTCTGTGCAAGATGGAG 388
QY      101  ValTyrHisGlyTyrValProTyrTyrPheGlyGlnGlyThrThrValThrValSerSerGly 120

```

```

DB      389  GTTTACCAAGGCTTACGTTCTTACTGGGGCCAAAGGACCAACAGTCCACGTTTCTTGGC 448
QY      121  GlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
DB      449  GGTGGCGGTTCTGTGGCGGTGGCTCCGGCGGTGGCGGTTTGACATCCAGCTAACCCAG 508
QY      141  SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaLaser 160
DB      509  TCTCACAATTCCTGTCACCTTCAGTGAAGACAGGGGTCAACATCACCTGCAAGGCCAGT 568
QY      161  GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB      569  CAGGATGTATTAATGCTGTGCTGGTATCAACAGAAACAGGACAAATCTCTAAACTT 628
QY      181  LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB      629  CTGATTTACTCGGCATCTCCGGGTACCTGAGGTCCCTTCTCGCTTCACTGGCAGTGGC 688
QY      201  SerGlyProAspPheThrPheThrIleSerSerValGlnIleAspLeuAlaValTyr 220
DB      689  TCTGGCGCGATTTCACTTTCACCATCAGCAGGTGTGACAGGCTGAAGACCTGGCAGTTAT 748
QY      221  PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
DB      749  TTCTGTGAGCAACATTTTCGTACTTCATTCACGTTCCGCTCGGGGACAAAATTGGAGATC 808
QY      241  LysAlaLeuGlnIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260
DB      809  AAAGCTCTAGAG-----CTGTTCTGGAA 832

```

Search completed: April 20, 2005, 10:44:55  
 Job time : 737.644 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005. Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 20, 2005, 10:19:14 ; Search time 224.44 Seconds  
(without alignments)  
3229.682 Million cell updates/sec

Title: US-09-596-774-7

Perfect score: 2345  
Sequence: 1 QIQVQSGPELKKKPGETVKI.....LSTAKDTYDALHQTIAPR 443

Scoring table:

BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 81813835 residues

Total number of hits satisfying chosen parameters: 240568

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+g2n.model -DEV=x1h  
-Q=/cgn2\_1/USPRO.epool/h/US09596774/runat\_20042005\_075443\_12533/app\_query.fasta\_1.1220  
-DB=Issued\_Patents\_NA -QPMT=fastcap -SUFFIX=aprx20.rml -MINMATCH=0.1 -LOOPCL=0  
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi  
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15  
-MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINTLEN=0 -MAXLEN=2000000000  
-USER=US09596774.QCGN\_1.1\_93 @runat\_20042005\_075443\_12533 -NCPU=6 -ICPU=3  
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPBLCK=100 -LONGLOG  
-DEV TIMEOUT=120 -WARN TIMEOUT=10 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -Fgapop=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

Issued\_Patents\_NA:\*  
1: /cgn2\_6/pctodata/1/ina/5A.COMB.seq:\*  
2: /cgn2\_6/pctodata/1/ina/5B.COMB.seq:\*  
3: /cgn2\_6/pctodata/1/ina/6A.COMB.seq:\*  
4: /cgn2\_6/pctodata/1/ina/6B.COMB.seq:\*  
5: /cgn2\_6/pctodata/1/ina/PTUS.COMB.seq:\*  
6: /cgn2\_6/pctodata/1/ina/backfiles.seq:\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1310.5	55.9	2012	1	US-08-235-838-13
2	1310.5	55.9	2012	2	US-08-465-473B-13
3	1308	55.8	1862	4	US-08-840-713-36
4	1308	55.8	1919	4	US-08-840-713-34
5	1305	55.7	1692	4	US-08-840-713-1
6	1303	55.6	2233	1	US-08-235-838-6
7	1303	55.6	2233	2	US-08-465-473B-6
8	1285.5	54.8	748	1	US-08-235-838-4
9	1285.5	54.8	748	2	US-08-465-473B-4
10	984.5	42.0	1869	2	US-08-356-786-15
11	981.5	41.9	909	1	US-08-133-804-1
12	981.5	41.9	909	1	US-08-461-184-7

13	981.5	41.9	909	1	US-08-463-675-7	Sequence 7, Appli
14	981.5	41.9	909	1	US-08-464-589-7	Sequence 7, Appli
15	981.5	41.9	909	1	US-08-461-838-1	Sequence 1, Appli
16	981.5	41.9	909	2	US-08-461-386-1	Sequence 1, Appli
17	903.5	38.5	723	1	US-07-988-430-93	Sequence 93, Appli
18	903.5	38.5	723	1	US-08-425-336-90	Sequence 90, Appli
19	903.5	38.5	723	1	US-08-488-113B-90	Sequence 90, Appli
20	903.5	38.5	723	1	US-08-477-484B-90	Sequence 90, Appli
21	903.5	38.5	723	2	US-08-646-186-90	Sequence 90, Appli
22	903.5	38.5	723	3	US-08-839-765-80	Sequence 90, Appli
23	903.5	38.5	723	3	US-09-136-389-90	Sequence 90, Appli
24	903.5	38.5	723	3	US-09-610-838-90	Sequence 90, Appli
25	903.5	38.5	723	4	US-09-711-485-90	Sequence 90, Appli
26	903.5	38.5	723	5	PCT-US92-09487-93	Sequence 93, Appli
27	892.5	38.1	1098	3	US-08-875-811-54	Sequence 54, Appli
28	890	38.0	1095	3	US-08-875-811-52	Sequence 52, Appli
29	889.5	37.9	1626	3	US-09-485-737B-84	Sequence 84, Appli
30	889.5	37.9	1626	4	US-10-071-485-84	Sequence 84, Appli
31	888.5	37.9	804	3	US-09-485-737B-1	Sequence 1, Appli
32	888.5	37.9	804	4	US-10-071-485-1	Sequence 1, Appli
33	888.5	37.9	2133	3	US-09-485-737B-89	Sequence 89, Appli
34	888.5	37.9	2133	4	US-10-071-485-89	Sequence 89, Appli
35	884.5	37.7	883	3	US-09-184-658-7	Sequence 7, Appli
36	884.5	37.7	883	4	US-09-504-262D-7	Sequence 7, Appli
37	878.5	37.5	1605	2	US-08-356-786-9	Sequence 9, Appli
38	866.5	37.0	786	3	US-08-635-928-11	Sequence 31, Appli
39	854	36.4	777	4	US-10-092-246-5	Sequence 5, Appli
40	854	36.4	777	4	US-10-092-246-7	Sequence 7, Appli
41	854	36.4	777	4	US-10-096-246A-5	Sequence 5, Appli
42	854	36.4	777	4	US-10-096-246A-7	Sequence 7, Appli
43	848	36.2	702	3	US-09-485-737B-92	Sequence 92, Appli
44	848	36.2	702	4	US-10-071-485-92	Sequence 92, Appli
45	846	36.1	777	4	US-10-092-246-6	Sequence 6, Appli

#### ALIGNMENTS

RESULT 1  
US-08-235-838-13  
Sequence 13, Application US/08235838  
Patent No. 5571894  
GENERAL INFORMATION:  
APPLICANT: Wels, Manfred S.  
APPLICANT: Hynes, Nancy E.  
APPLICANT: Harwerth, Ina-Maria  
APPLICANT: Groner, Bernd  
APPLICANT: Hardman, No. 5571894man  
TITLE OF INVENTION: Recombinant Antibodies Specific for a  
TITLE OF INVENTION: Growth Factor Receptor  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10532  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235,838  
FILING DATE: TBA  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/828,832  
FILING DATE: 31-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 91-810079.3  
FILING DATE: 05-FEB-1991

ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: 4-18518/A/CLP/CONT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8614  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEO ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2012 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mouse/Pseudomonas aeruginosa  
INDIVIDUAL ISOLATE: E. coli  
IMMEDIATE SOURCE:  
CLONE: pMW215-5  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 1..1911 /note= "64-87 FLAG peptide and  
OTHER INFORMATION: enterokinase cleavage site; 97-453 heavy chain  
OTHER INFORMATION: variable domain;454-498 15 aa linker sequence;  
OTHER INFORMATION: 499-822 FRP5 light chain variable domain  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 1..63  
OTHER INFORMATION: /note= "ompA signal peptide"  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 94..1911  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 1912..2012  
OTHER INFORMATION: /function= "3'non-coding region of  
OTHER INFORMATION: the exotoxin A gene"  
FEATURE:  
NAME/KEY: misc\_feature  
LOCATION: 826..1911  
OTHER INFORMATION: /note= "Exotoxin A gene coding  
OTHER INFORMATION: region (coding for amino acids 252 to 613 of the  
OTHER INFORMATION: mature exotoxin A)"  
US-08-235-838-13  
Alignment Scores:  
Pred. No.: 9,37e-117 Length: 2012  
Score: 1310.50 Matches: 279  
Percent Similarity: 68.51% Conservative: 19  
Best Local Similarity: 64.14% Mismatches: 43  
Query Match: 55.88% Indels: 95  
DB: 1 Gaps: 16  
US-09-596-774-7 (1-443) x US-08-235-838-13 (1-2012)  
QY 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLeuValSerProGlyGlnThrValValSile 20  
DB 97 CAGGTACAACTGCAGCGCTGACCTGACCTGAAGAAAGCTGAGAGACAGCAAGATC 156  
QY 21 SerCysLeuValSerGlyTyrProPheThrAsnTyrGlyMetAsnThrValValSileGlnAla 40  
DB 157 TCCGCAAGGCTCTGGGTATCCTTCAACAACCTATGATGAATGAATGGGTGAAGCAAGCT 216  
QY 41 ProGlyGlnGlyLeuValSerGlyTyrProPheThrAsnTyrGlyMetAsnThrValValSileGlnAla 60  
DB 217 CCAGGACAGAGGTAAAGTGAATGAGCTGATTAACACTCACTGAGAGTCAACATTT 276  
QY 61 AlaAspAspPheLeuSerGlyTyrArgPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80  
DB 277 GCTGATGACTTCAAGGAGCGGTGACTTCTTTGGAAACCTCTGCAACACTGCTAT 336

QY 81 LeuGlnIleAsnLeuValSerGlyAspMetAlaThrTyrPheCysAlaAaGTPGlu 100  
DB 337 TTGAGATCAACAACCTCAAAAGTGAAGATGGCTACATATTTCTGTGCAAGTGGAG 396  
QY 101 ValTyrIleGlyTyrValProTyrTyrGlnGlyThrThrValThrValSerSerGly 120  
DB 397 GTTACCAAGCGCTACGTTCTTACTGGGACCAAGGACCAAGCTACCGTTCTCTGGC 456  
QY 121 GlyIleGlySerGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 140  
DB 457 GGTGGCGTTCTGGTGGCGGTGGCTCCGCGGTGGCGGTCTTGACATCCAGCTGACCCAG 516  
QY 141 SerHisIlePheLeuSerThrThrSerValGlyAspArgValSerIleThrCysLeuAlaSer 160  
DB 517 TCTCAAAATTCCTGCTCCACTTCAGTAGAGAGAGAGGTCAGATCACTTCAGAGCGCAGT 576  
QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLeuSerProGlyGlnSerProGlyLeu 180  
DB 577 CAGGATGTGTAATGCTGTGCTGCTGATCAACAGAAACCAAGACATCTCTAATCT 636  
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
DB 637 CTGATTACTCGGCATCTCCCGGTACCTGAGATCCCTTCTGCTTCACTGGCAGTGGC 696  
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGlnLeuAlaValTyr 220  
DB 697 TCTGGCGCGATTTCACTTTCACCATCAAGTGTGAGGCTGAAACCTGGCAGTTTAT 756  
QY 221 PheCysGlnGlnIlePheArgThrProPheThrPheGlySerGlyThrIleLeuGlnIle 240  
DB 757 TTCTGTACAGAACATTTTGTACTTCATTCAGCTTGGCTGGGAGCAAAATTTGAGATC 816  
QY 241 LysAlaLeuGlnIle-----IleSerAsnSerValMetTyrPheSerSerValValPro 257  
DB 817 AAAGCTCTAGAGGGCGCGCAGCTGGCGCGCTGACCGCCAGCAGCGCTGCCACTGGCG 876  
QY 258 ValLeuGlnIleValAsnSerThrThrThrIle----- 268  
DB 877 CTG-----GAGACTTTCACCGCTCATCGCATCGCCAGCGCGCGCTGGGAACA 921  
QY 269 -----ProValLeuArgThrProSerProValIleProThrGlyThr 282  
DB 922 CTGAGAGAGTGGCGCTATCGGTGAGCGCGCTGCT--CGCCTCTA---CTGGCGCGCG 977  
QY 283 Ser-----GlnProGlnArgProGlyAspCysArgProArgIleSerValIleGlyThr 300  
DB 978 ACTGTCATGGAACCAAGTCGACGACGAGTGAATCCGCAACGCCCTGGCCAGCGCGAGCGG 1037  
QY 301 GlyLeuAspPheLeuGlnAspProGlyLeuGlyTyrLeuLeuAspGlyIleLeuPheIle 320  
DB 1038 CGCGCACTGGCGGAGCAACCAATCG----- 1061  
QY 321 TyrGlyValIleIleThrAlaLeuTyrIleuArgAlaIlePheSerArgSerAlaGlnThr 340  
DB 1062 -----CGAGCACCGCGAGCA 1076  
QY 341 AlaAlaAsnLeuGlnAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlnValArgGln 360  
DB 1077 GGCCTCTGCGCTGACCT-----GGCGCGCGCGCA 1109  
QY 361 GlyTyrAspValLeuGlnIleValSerGlyAlaArgAspProGlyMetGlyGlyIleGln 380  
DB 1110 GAG-----CGAGCGCTTCTGCTCG-----CGAGGCGCACCGGCA 1142  
QY 381 ArgArgArgAsnProGlnGlnIleValTyrAsnAlaLeuGlnIleAspIleMetAlaGln 400  
DB 1143 CGACGA----- 1148  
QY 401 AlaTyrSerGlnIleGlyThrIleGlyValIleArgArgGlyValys 415  
DB 1149 -----GGCGCGCGCGCGCAACCGCGAGCTGTGAG 1178



```

RESULT 2
US-08-465-473B-13
Sequence 13, Application US/08465473B
Patent No. 5939531
GENERAL INFORMATION:
APPLICANT: Weis, Winfried S.
APPLICANT: Hynes, Nancy E.
APPLICANT: Harwerth, Ina-Maria
APPLICANT: Groner, Bernd
APPLICANT: Hardman, No. 5939531man
APPLICANT: Zwickl, Markus
TITLE OF INVENTION: Recombinant Antibodies Specific for a
TITLE OF INVENTION: Growth Factor Receptor
NUMBER OF SEQUENCES: 34
CORRESPONDENCE ADDRESS:
ADDRESSEE: NOVARTIS Corporation
STREET: 564 Morris Avenue
City: Summit
STATE: New Jersey
COUNTRY: USA
ZIP: 07901-6940
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/465,473B
FILING DATE: 5 June 1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 07/828,832
FILING DATE: 31-JAN-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: GB 91-810079.3
FILING DATE: 05-FEB-1991
ATTORNEY/AGENT INFORMATION:
NAME: Pfeiffer, Heena J.
REGISTRATION NUMBER: 22,640
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2
TELECOMMUNICATION INFORMATION:
TELEPHONE: (908)522 6940
TELEFAX: (908)522 6955
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 2012 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA
HYPOTHETICAL: NO
ANTI-SENSE: NO
ORIGINAL SOURCE:
ORGANISM: Mouse/Pseudomonas aeruginosa
INDIVIDUAL ISOLATE: E. coli
IMMEDIATE SOURCE:
CLONE: pMW215-5
FEATURE:
NAME/KEY: CDS
LOCATION: 1..1911 /note= "64-87 FLAG peptide and
OTHER INFORMATION: enterokinase cleavage site; 97-453 heavy chain;
OTHER INFORMATION: variable domain;454-498 15 aa linker sequence;
OTHER INFORMATION: 499-822 FRP5 light chain variable domain
FEATURE:
NAME/KEY: sig_peptide
LOCATION: 1..63
OTHER INFORMATION: /note= "ompA signal peptide"
FEATURE:
NAME/KEY: mat_peptide
LOCATION: 94..1911
FEATURE:
NAME/KEY: 3'UTR

```

[illegible]



GENERAL INFORMATION:  
APPLICANT: WELS, Winfried, Dr.  
APPLICANT: FOYMINA, Jesus  
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby  
City: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/840,713  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kites, Monica Chin  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: 1614-7014  
TELEPHONE: (202) 638 - 5000  
TELEFAX: (202) 638 - 4810  
INFORMATION FOR SEQ ID NO: 34:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1919 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 64..1908  
US-08-840-713-34

Alignment Scores:  
Pred. No.: 1,52e-116 Length: 1919  
Score: 1308.00 Matches: 245  
Percent Similarity: 97.64% Conservative: 3  
Best Local Similarity: 96.46% Mismatches: 6  
Query Match: 55.78% Indels: 0  
DB: 4 Gaps: 0

US-09-596-774-7 (1-443) x US-08-840-713-34 (1-1919)

QY 1 Glnlleglnleuvalglnsenglyprogluleuylsypbroglygluthrvalylle 20  
DB 1147 CAGGTACACTGCAGCAGTCTGCAGCTGAACTGAAGAGCCTGGAGAGACGTAACATC 1206  
QY 21 SerCyslvsalaserglytyrprophethrasnlyrglymetcantspvalylslnala 40  
DB 1207 TCTGCGAAGGCTTGCGGTATCTTTCACAAACTATGGAATGAATGAGTGAACAGCT 1266  
QY 41 ProglYglnlYleuLYsTIpMeTgLYTIpIleasnthrSerthrglygluSerThrPhe 60  
DB 1267 CCAAGACAGGGTTTAAAGTGATGGCTGGATTAACACCTCACTGAGAGATCAACATTT 1326  
QY 61 AlaAspAspPheLYsgLYaTgPheAspPheSerleuGLuThrSerAlaAsnthrAlaTyr 80  
DB 1327 GCTGATGACTTCAGAGGACGGTTGACTTCTCTTGGAAACCTTGCCAAACACTGCCTAT 1386  
QY 81 LeuGLInlleaAsnleuLYsSerGLuAspMeTAlaThrTYrPheCYsAlaTgTIpGLu 100  
DB 1387 TTGCAAGATCAACAACCTCAAAAGGAAGACATGCTCATATTTCTGTGCAAGTGGAG 1446  
QY 101 ValTYrHisGLYTYrVal1ProTYrTIpGLYglnlYThrThrVal1ThrValSerSerGLY 120  
DB 1447 GTTACACAGCGCTACGTTCTTACTGGGGCCAAAGGACCAAGGTCACCGTTTCTCTGGC 1506

QY 121 GLYGLYGLYserGLYGLYGLYserGLYGLYGLYserAspIleGlnLeuThrGln 140  
DB 1507 GGTGGCGGTTCTGTGGCGGTGGCTCCGGCGGTGGCGGTTCTGACATCCAGTACCCAG 1566  
QY 141 SerHisAspPheLeuSerThrservalgLYAspArgValserIlethrCYsLYsAlaSer 160  
DB 1567 TCTCACAAATTCCTGTGCATCTCAGTGAAGACAGGGTCAAGATCACTCCAAAGCCAGT 1626  
QY 161 GlnAspValTYrAsnAlaValaTTPYrGLInLYsProGLYglnSerProLYsLeu 180  
DB 1627 CAGATGTGTATATGCTGTGGCTGGTATCAACAGAAACAGACAAATCTCTAAACTT 1686  
QY 181 LeuIleTYrSerAlaSerSerArgTYrThGLYVal1ProSerAspPheThrgLYserGLY 200  
DB 1687 CTGATTTACTCGGATCTCCCGGTACACTGAGATCCCTTCTCGCTTCACTGCGAGTGC 1746  
QY 201 SerGLYProAspPheThrPheThrIleSerSerValGlnaGLuAspLeuAlaValTYr 220  
DB 1747 TCTGGCGCGGATTTCACTTCAACATCAGAGTGTGAGGCTGAAGACCTGGCAGTTTAT 1806  
QY 221 PheCYglnGLInHisPheArgThrProPheThrPheGLYserGLYThrLYsLeuGLuIle 240  
DB 1807 TTTGTGACGAACATTTTCGTACTCATTCACTGCGCTCGGGGACAAATTTGAGATC 1866  
QY 241 LYsAlaLeuGLuIleSerAsnSerValMetTYrPheSerSer 254  
DB 1867 AAGCTTAAAGATCTCTCGAGTGAAGAAATTTTACACC 1908

RESULT 5  
US-08-840-713-1  
Sequence 1, Application US/08840713  
Patent No. 6498233  
GENERAL INFORMATION:  
APPLICANT: WELS, Winfried, Dr.  
APPLICANT: FOYMINA, Jesus  
TITLE OF INVENTION: NUCLEIC ACID TRANSFER SYSTEM  
NUMBER OF SEQUENCES: 58  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: Nikaido, Marmelstein, Murray & Oram LLP  
STREET: 655 15th St., N.W., Suite 330 - G St. Lobby  
City: Washington  
STATE: D.C.  
COUNTRY: USA  
ZIP: 20005-5701  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.30  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/840,713  
FILING DATE: 25-APR-1997  
CLASSIFICATION: 514  
ATTORNEY/AGENT INFORMATION:  
NAME: Kites, Monica Chin  
REGISTRATION NUMBER: 36,105  
REFERENCE/DOCKET NUMBER: 1614-7014  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (202) 638 - 5000  
TELEFAX: (202) 638 - 4810  
INFORMATION FOR SEQ ID NO: 1:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 1692 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
IMMEDIATE SOURCE:  
CLONE: pMF46-5  
FEATURE:  
NAME/KEY: sig peptide  
LOCATION: 1..63



REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (919) 541-8614  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2233 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mouse and E. coli  
INDIVIDUAL ISOLATE: E. coli  
IMMEDIATE SOURCE:  
CLONE: pmw616  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 23..2155  
OTHER INFORMATION: /note= "89-445 FRP5 heavy chain  
var. domain; 446-490 is a linker sequence; 491-814  
OTHER INFORMATION: FRP5 light chain var. domain; 815-2155 coding region  
OTHER INFORMATION: of phoA  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..22  
OTHER INFORMATION: /function= "ompa 5'non-coding  
OTHER INFORMATION: region"  
FEATURE:  
NAME/KEY: sig peptide  
LOCATION: 23..85  
OTHER INFORMATION: /note= "ompa signal peptide"  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 2156..2233  
OTHER INFORMATION: /function= "phoA 3' non-coding  
OTHER INFORMATION: region"  
FEATURE:  
NAME/KEY: mat peptide  
LOCATION: 86..2155  
US-08-235-838-6  
Alignment Scores:  
Pred. No.: 5.8e-116 Length: 2233  
Score: 1303.00 Matches: 245  
Percent Similarity: 95.00% Conservative: 2  
Best Local Similarity: 94.23% Mismatches: 1  
Query Match: 55.57% Indels: 12  
DB: 1 Gaps: 1  
US-09-596-774-7 (1-443) x US-08-235-838-6 (1-2233)  
QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysLeu 20  
DB 89 CAGGTACCACTGCGACGCTGACCTGAACTGAAGAGCTGGAGACGTCAGATC 148  
QY 21 SerCysValAspAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIleGlnAla 40  
DB 149 TCCGCAAGGCTCTGGGTATCCTTTCACAACTATGSAATGAACTGGGTAAAGCAAGCT 208  
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrPheLeuThrSerThrGlyGluSerThrPhe 60  
DB 209 CCAGAGACGGGCTTTAAAGTGAATGGCTGAGTTAAACCTCAGCTGAGAGCTCAACATTT 268  
QY 61 AlaAspApphLeuValGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80  
DB 269 GCTATGACTTCAAGGAGCGGTGACTTCTCTTGGAAACCTCTGCAACAGCTGCTAT 328  
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTrpGlu 100  
DB 329 TTGCAATCAACCACTCAAAAGTGAAGACATGGCTATATTTCTGTGCAAGATGGAG 388

QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrThrValThrValSerSerGly 120  
DB 389 GTTACCAACGCGTACCTCTTCTTACTGGGCGCAAGGACACAGGTCACCTTCTCTGGC 448  
QY 121 GlyGlySerGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerGlyGlyGly 140  
DB 449 GGTGGCGTTCTGGTGGCGGTGGCTCCGCGGTGGCGGTGGCTGGCAATCCAGTACCAG 508  
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSer 160  
DB 509 TCTCAAAATTCCTGTCCTCCTCAGTAGAGACAGGCTCAGCATCCTGAAAGCCAGT 568  
QY 161 GlnAspValTyrAsnAlaValAlaTrpTyrGlnGlnLysProGlyGlnSerProLysLeu 180  
DB 569 CAGATGTGTATATGCTGTGCTGCTGATATCAACAGAAACAGACATCTCTAACTT 628  
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
DB 629 CTGATTTACTCGGCATCTCCCGGTACATGAGAGTCCCTTCTGCTTCACTGGCAGTGGC 688  
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220  
DB 689 TCTGGCGCGGATTTCTTCTTCACTCAGCAGAGTGTGAGGCTGAAGACTGGCAGTTAT 748  
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrThrLysLeuGlu 240  
DB 749 TTCTGACGACCACTTTCTGACTCTCATTCACGCTTCGGCTCGGGACAAATTTGAGATC 808  
QY 241 LysAlaLeuGluIleSerAsnSerValMetTyrPheSerSerValProValLeuGln 260  
DB 809 AAAGCTTAGG-----CCTGTTCTGGAA 832  
RESULT 7  
US-08-465-473B-6  
Sequence 6, Application US/08465473B  
Patent No. 5939531  
GENERAL INFORMATION:  
APPLICANT: Wels, Winfried S.  
APPLICANT: Hynes, Nancy E.  
APPLICANT: Harwerth, Ina-Maria  
APPLICANT: Groner, Bernd  
APPLICANT: Hardman, No. 5939531man  
APPLICANT: Zwickl, Markus  
TITLE OF INVENTION: Recombinant Antibodies Specific for a  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NOVARTIS Corporation  
STREET: 564 Morris Avenue  
CITY: Summit  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07901-6940  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent in Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,473B  
FILING DATE: 5 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/828,832  
FILING DATE: 31-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 91-810079.3  
FILING DATE: 05-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Pfeiffer, Heena J.  
REGISTRATION NUMBER: 22,640  
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2

TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908)522.6940  
TELEFAX: (908)522.6955  
INFORMATION FOR SEQ ID NO: 6:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 2233 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE:  
ORGANISM: Mouse and E. coli  
INDIVIDUAL ISOLATE: E. coli  
IMMEDIATE SOURCE:  
CLONE: pMW616  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 23..2155  
OTHER INFORMATION: /note= "89-445 FRP5 heavy chain  
var. domain; 446-490 15 aa linker sequence; 491-814  
FRP5 light chain var. domain; 815-2155 coding region  
of phoA  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: 5'UTR  
LOCATION: 1..22  
OTHER INFORMATION: /function= "ompa 5'non-coding  
region"  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: sig\_peptide  
LOCATION: 23..85  
OTHER INFORMATION: /note= "ompa signal peptide"  
FEATURE:  
NAME/KEY: 3'UTR  
LOCATION: 2156..2233  
OTHER INFORMATION: /function= "phoA 3' non-coding  
region"  
OTHER INFORMATION:  
FEATURE:  
NAME/KEY: mat\_peptide  
LOCATION: 86..2155  
US-08-465-4738-6  
Alignment Scores:  
Pred. No.: 5.8e-116 Length: 2233  
Score: 1303.00 Matches: 245  
Percent Similarity: 95.00% Conservative: 2  
Best Local Similarity: 94.23% Mismatches: 1  
Query Match: 55.57% Indels: 12  
DB: 2 Gaps: 1  
US-09-596-774-7 (1-443) x US-08-465-4738-6 (1-2233)  
QY 1 Glnlleglnleuvalglnserglyprogluleuylsyproglyglulthrvallystle 20  
DB 89 CAGGTAACTGACGACGTGTGACCTGAACTGAAGAGCCTGGAGAGACATCAGATC 148  
QY 21 SerCysylsyalaserglytyrprophethrasnthyrglywelsantpvallysglnala 40  
DB 149 TCCGCGAAGGCTCTGGGTATCTTTTCAAACTATGAAATGAACTGGGTGAGCAGGCT 208  
QY 41 Proglylgnllyleuylstipmetglytrpilesanthserthrglygluserthp 60  
DB 209 CCAGGACAGGCTTTAAAGTGAATGGGCTGATTAACCTCCACCTGAGAGTCAACTTT 268  
QY 61 Alaasaphselysglyargphesphserleuglulthseralaaenthralaaty 80  
DB 269 GCTGATGACTTCAAGGACGCTTGACTTCTTTGGAAACCTCTGCAACACTGCTAT 328  
QY 81 LeuGlnlleasnleuylserglyaspmetAlathrtyrphecysalaaetpplu 100  
DB 329 TTGCAATCAACAACTCAAAAGTGAAGACATGGCTACATATTTCTGTGCAAGTGGAG 388

QY 101 Valtyrhsglytyrvalprotyrtmglyglnlgythrthrvalthrserserly 120  
DB 389 GTTACACAGCTACGTTCTTACTGGGGCCAAAGGACACGGTCACTTCTCTGGC 448  
QY 121 Glylgllyserglyglyglyserglyglylgllylgllyseraspilleglnleu 140  
DB 449 GGTGGCGGCTTCTGGTGGCGGTGGGTCCGCGGGGGGCTTCAACATCCAGTCA 508  
QY 141 Serhslyspheleuserthrservalglyaspharyalsertlethrcyslylaser 160  
DB 509 TCTCAAAATTCCTGTCACCTCAGTAGACAGGCTCAGTCACTGCAAGGCGAGT 568  
QY 161 Glaspvaltyrasnalaalattprtyrglnllyserproglylgnserprolyleu 180  
DB 569 CAGGATGTATATGCTGTGCTGGTATCAAGAAACAGACAGACATCTCTAACTT 628  
QY 181 leuiletyseralaserargtyrthrglyvalproserarphethrglysergly 200  
DB 629 CTGATTTACTGGCATCTCCGGTACACTGAGATCCCTTCTGCTTCACTGGCAGTGGC 688  
QY 201 Serglyproasphethrphethrilleserervalglnlaqluaspleualaavaly 220  
DB 689 TCTGGCGCGATTTCACTTCAACATCAGCAGGTGACAGGCTGAAAGCCTGGCAGTTAT 748  
QY 221 Phcysglnlgnhispheargthrprophethrphcglyserglythrllysleuglu 240  
DB 749 TTCTGTCAGCAACATTTCTTCACTTCACTTCACTTGGGCGCAAAATTTGGAGATC 808  
QY 241 LysalaleuGlnlleaseraservalmettyrpheserervalvalprovalleugln 260  
DB 809 AAGCTCTAGAG-----CCTGTTCTGAA 832

RESULT 8  
US-08-235-838-4  
Sequence 4, Application US/08235838  
Patent No. 5571894  
GENERAL INFORMATION:  
APPLICANT: Weis, Winfried S.  
APPLICANT: Hynes, Nancy E.  
APPLICANT: Harwerth, Ina-Maria  
APPLICANT: Groner, Bernd  
APPLICANT: Hardman, No. 5571894man  
APPLICANT: Zwickl, Markus  
TITLE OF INVENTION: Recombinant Antibodies Specific for a  
NUMBER OF SEQUENCES: 16  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: CIBA-GEIGY Corporation  
STREET: 7 Skyline Drive  
CITY: Hawthorne  
STATE: New York  
COUNTRY: USA  
ZIP: 10512  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patent Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/235, 838  
FILING DATE: TBA  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/828, 832  
FILING DATE: 31-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 91-810079.3  
FILING DATE: 05-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Elmer, James Scott  
REGISTRATION NUMBER: 36,129  
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT  
TELECOMMUNICATION INFORMATION:

TELEPHONE: (919) 541-8614  
TELEFAX: (919) 541-8689  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 748 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Mouse  
INDIVIDUAL ISOLATE: E. coli  
IMMEDIATE SOURCE:  
CLONE: pMW52  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 6..728  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: 1..701 /note= "1-8 synthetic spacer;9-365  
OTHER INFORMATION: FRP5 heavy chain var. domain;99-113 CDR1H;156-206  
OTHER INFORMATION: CDR2H;303-332 CDR3H;366-410 15 aa linker seq;411-728  
OTHER INFORMATION: FRP5 light chain var dom;480-512 CDR1L;558-578 CDR2L;  
OTHER INFORMATION: 675-701 CDR3L  
US-08-235-838-4

Alignment Scores:  
Pred. No.: 5.7e-115 Length: 748  
Score: 1285.50 Matches: 241  
Percent Similarity: 98.37% Conservative: 1  
Best Local Similarity: 97.87% Mismatch: 3  
Query Match: 54.82% Indels: 3  
DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-235-838-4 (1-748)

QY 1 Glnlleglnleuvalglnserglyprogluleuylsyproglygluthvallyalle 20  
DB 9 CAGGTACACGCGCTTAAAGTGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCT 68  
QY 21 SerCySylsAlaserglytyrProPhenThraenTyrglyMeCaantTTPvallyeGlnAla 40  
DB 69 TCCTGCAAGGCTTGGGTATCTCTTTCACAACTATGGAAGAACTGGGTGAAGGCT 128  
QY 41 ProglynglyleuylsyrpmetglytrpIleasnThrserrhnglygluSerThrPhe 60  
DB 129 CCAGGACAGGCTTAAAGTGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCT 188  
QY 61 AlaAspPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 80  
DB 189 GCTGATGACTTCAAGGCGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCTTGAAGGCT 248  
QY 81 leuGlnIleAsnleuvalglnserglyprogluleuylsyproglygluthvallyalle 100  
DB 249 TTGAGAGATCAACCACTTAAAGTGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCT 308  
QY 101 ValTyRhlsglytyrvalprolytrpIleasnThrserrhnglygluSerThrPhe 120  
DB 309 GTTACACAGGCTTAAAGTGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCTGGAAGGCT 368  
QY 121 Glyglyserglyglyglyglyglyglyglyglyglyglyglyglyglyglyglyglygly 140  
DB 369 GGTGGCGGCTTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 428  
QY 141 SerHisylsPheleuserThrservalgIyAspArgvalserIleThrCyAlaAlaser 160  
DB 429 TCTACAAATTCCTGCTCAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT 488  
QY 161 GlnAspValTyRAsnAlaValAlaIlePlyrGlnIleuylsyproglygluSerProlyleu 180

DB 489 CAGGATGTGATTAATGCTGTGCTGGTATCAACAGAAACAGAGCAATCTCTAACTT 548  
QY 181 leuIleTySerAlaserglytyrThrglyvalProserArgPheThrglySerGly 200  
DB 549 CTGATTTACTCGGATCTCTCCGATACAGTGAATCTCTCTCTCTCTCTCTCTCTCTCTCT 608  
QY 201 SerGlyProAspPhePhePhePhePhePhePhePhePhePhePhePhePhePhePhePhe 220  
DB 609 TCTGGCGGATTTCACTTTCACCATCAGAGTGTGCAAGGCTGAAGACTGGCAGTTTAT 668  
QY 221 PheCySylsAlaserglytyrProPhenThraenTyrglyMeCaantTTPvallyeGlnAla 239  
DB 669 TTCTGAGCAAACTTTTCGATCTCACTCACTCACTCACTCACTCACTCACTCACTCACTCA 728  
QY 240 -----IlelysaIaleu 243  
DB 729 TAGCTGATCAAGCTCTTA 746  
RESULT 9  
US-08-465-473B-4  
Sequence 4, Application US/08465473B  
Parent No. 5939531  
GENERAL INFORMATION:  
APPLICANT: Wels, Winfried S.  
APPLICANT: Hynes, Nancy E.  
APPLICANT: Harwerth, Ina-Maria  
APPLICANT: Groner, Bernd  
APPLICANT: Hardman, No. 5939531man  
APPLICANT: Zwickl, Markus  
TITLE OF INVENTION: Recombinant Antibodies Specific for a  
TITLE OF INVENTION: Growth Factor Receptor  
NUMBER OF SEQUENCES: 34  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: NOVARTIS Corporation  
STREET: 564 Morris Avenue  
CITY: Summit  
STATE: New Jersey  
COUNTRY: USA  
ZIP: 07901-6940  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/465,473B  
FILING DATE: 5 June 1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: US 07/828,832  
FILING DATE: 31-JAN-1992  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: GB 91-810079.3  
FILING DATE: 05-FEB-1991  
ATTORNEY/AGENT INFORMATION:  
NAME: Pfeiffer, Heena J.  
REGISTRATION NUMBER: 22,640  
REFERENCE/DOCKET NUMBER: 4-18518/A/CIP/CONT2  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (908) 522 6940  
TELEFAX: (908) 522 6955  
INFORMATION FOR SEQ ID NO: 4:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 748 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: DNA  
HYPOTHEICAL: NO  
ANTI-SENSE: NO  
ORIGINAL SOURCE: Mouse  
INDIVIDUAL ISOLATE: E. coli





Qy	41	ProGlyGlnGlyLeuLeuYTrpMetGlyTrpIleAsnThrSerThiGlySerThrPhe	60
Db	127	CCAGGAAGGGGCTTAAAGTGGAGGCGGTGGATAAACCAACACCTGGAGAGCCAACTATT	186
Qy	61	AlaAspAspPheLeuGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTrp	80
Db	187	GCTGAAGAGTTCAAGGGAACGGTTGCCCTTCTTTGGAAACCTCTGGCAGACCTGCATT	246
Qy	81	LeuGlnIleAsnAsnLeuLeuSerGluAspMetAlaThrTrpPheCysAlaArgTrpGlu	100
Db	247	TTGCAGATCAACCACTCAAAATATGAGCACGGCTACATATTCTGTGAAGGCAATTT	306
Qy	101	ValTrpHisGlyTrpValProTrpTrpGlyGlnGlyThrThrValThrValSerSerGly	120
Db	307	ATTACCTTAACGGGGGTGGTCTTAAGTGGGGCCAAAGGAACTCTGGTCACTGTCTCTGA	363
Qy	121	GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln	140
Db	364	TCGAGCTCTCCGATCTTCAATCTAGCGGTTCCAGCTGAGGGAATATGTCATAGACCG	423
Qy	141	SerHisLeuPheLeuSerThrSerValGlyAspArgValSerIleThrCysAlaSer	160
Db	424	TCTCTTAATTCATCTCAACGTCAAGTGGGAGACAGGGTCAAGATCTCCGCAAGCCAGT	483
Qy	161	GlnAspValTrpAsnAlaValAlaThrPyrGlnGlnLeuProGlyGlnSerProLeuLeu	180
Db	484	CAGAGTGGAGTACTGCTGTAGCCCTGGTATCAACAAAAACAGGCGCAATCTCTTAAC	543
Qy	181	LeuIleTrpSerAlaSerSerArgTrpThrGlyValProSerArgPheThrGlySerGly	200
Db	544	CTGATTTACTGGACATCCACCCGGACACACTGGAATCTCGATCCGTTACACAGCAGTGA	603
Qy	201	SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTrp	220
Db	604	TCTGGGACAGATTATTACTCTCACCATCAGACAGTGTGAGGCTGGAAGACTGGCACCT	663
Qy	221	PheCysGlnGlnHisPheArgThrTrpPheThrPheGlySerGlyThrTrpLeuGluIle	240
Db	664	TACTGTCCGACCACTATTATAGAGTGGCCACAGTTTCGGAAGGGGACCAAGCTGGAGATA	723
Qy	241	LeuAlaLeuGlnIleSerAsnSerValMetTrpPheSerSerValAlaProValLeuGln	260
Db	724	AAACGGGCTGATGCTGCACCAACTGTA-----TCAATCTTCCACCATTCAGT	771
Qy	261	LeuValAsnSerThrThrThrLeuProValLeuArgThrProSerProValHis--Pro	279
Db	772	GAGCAGTTTGAAGGGGGGAGCGCTGGCGCGGCTGAAGACGACACAGGCTTGCCACCGCG	831
Qy	280	ThrGlyThrSerGlnProGlnArgProGluAspCysArgProArgGlySerValIleGly	299
Db	832	CTGGAGACTTTCACCCGTCATCGC-----CAGCCCGCGGCG-----	867
Qy	300	ThrGlyLeuAspPheLeuGluAspProLeuLeuCys--TyrLeuLeuAspGlyIleLeu	318
Db	868	-----TGGAAACAACTGAGGACGTGGGCTATCCGGTGCAGCGGCTGTC	912
Qy	319	PheIleTrpGlyValIleIleThrAlaLeuTrpLeuArgAlaIlePheSer-----	335
Db	913	-----GCCCTTACCTGGCGCGCGCGGTGTCGTGGAAACG	948
Qy	336	-----ArgSerAlaGluThrAla-----	341
Db	949	GTCGACCACTGATCCGCAACCGCTGGCGCACCCCGGACGCGCGGCGACCTGGGCGAA	1000
Qy	342	AlaAsnLeuGlnAspProAsnGlnLeuTrpArgGlnLeuAsnLeuGlyValArgTrpGlnGlu	361
Db	1009	GGGATCCCGAGACACCGGAGAGGCGCGCTCGGCCCTTACCTT--GGCCGCGCGGAGAG	1067
Qy	362	TyrAspValLeuGlnIleLeuIleValArgAlaArgAspProGluMetGlyGlyIleGlnGln	381
Db	1068	-----CGAGCGCTTCGCTCG-----GCAGGAGCACCGGACAGCA	1100
Qy	382	ArgAspAsnProGlnGluGlyValTrpAsnAlaLeuGlnLeuAspIleMetAlaGluAla	401

```

Db      1101 CGA-----1101
Cy      402 TyrSerGluIleGlyThrLysGlyGluArgArgGlyLys 415
Db      1104 -----GGCCGGCGCGCGCAACGCCGAGTGTGTGAG 1133

RESULT 11
US-08-133-804-1
; Sequence 1, Application US/08133804
; Patent No. 5534254
; GENERAL INFORMATION:
; APPLICANT: Huston, James S.
; APPLICANT: Oppermann, Hermann
; APPLICANT: Houston, L. L.
; APPLICANT: Ring, David B.
; TITLE OF INVENTION: Biosynthetic Binding Proteins For
; TITLE OF INVENTION: Imaging
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department
; STREET: Exchange Place, 53 State Street
; CITY: Boston
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02109
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/133,804
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Kelley, Robin D.
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: 2054/22
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617-248-7477
; TELEFAX: 617-248-7100
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 909 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..752
; OTHER INFORMATION: /product= "741F8 gFv' C-terminal
; US-08-133-804-1.

Alignment Scores:
Pred. No.: 1.84e-85 Length: 909
Score: 981.50 Matches: 185
Percent Similarity: 85.89% Conservative: 22
Best Local Similarity: 76.76% Mismatch: 33
Query Match: 41.86% Indels: 1
DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-133-804-1 (1-909)
Cy      1 GlnTlEgLnLeuValGlnSerGlyProGluLeuLysLysProGlyGluThrValLysIle 20
Db      9 GAGATTCACAAATGTGGCGAGCTCTGACCTGAGCTGAAGAAGCCTGGAGAGACAGTAAAGTC 68
Cy      21 SerCysValValAspSerGlyTyrProPheThrAsnTyrGlyMetAsnTPValLysGlnAla 40
Db      69 TCTCTGCAAGGCTTTGTGGGTATACCTTACACAAACTATGCAATGAATGAATCTGGTGAACAGGCT 128

```

```

QY      41      ProGlyGlnGlyLeuLeuTyrPheMetGlyTyrPheLeuThrSerThrGlyGluSerThrPhe 60
      129      CCAGGAAGGCTTTAAAGTGGATGGCTGGATAAACCAACCACTGGAGAGCCACATAT 188
QY      61      AlaAspAspPheLeuGlyValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
      189      GCTGAAGAGTTCAAGGAGCGGTTTGCCCTTCTTGGAAACCTCGCAGACAGCTGCTAT 248
QY      81      LeuGlnIleAsnAsnLeuLeuSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
      249      TTGCAGATCAACAACCTCAAAATAGAGACACGGCTACATATTTCTGTGGAAGCAATT 308
QY      101      ValTyrHisGlyTyrValProTyrTyrPheGlyGlnGlyThrThrValThrValSerSerGly 120
      309      ATTACCTACGCGGGTTTGCTACCTGGGCGCAAGGAGTCTGCTGCTGCTGCA--- 365
QY      121      GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
      366      TCGAGCTCCTCCGATCTTCATCTAGCGGTTCCAGCTCGACGATATCGTCATGACCCAG 425
QY      141      SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSer 160
      426      TCTCTTAATTCATGCTCCACGCTCAGTGGAGACAGGCTCAGCATCTCTGCAAGGCCAGT 485
QY      161      GlnAspValTyrAsnAlaValAlaTyrPheGlnGlnLysProGlyGlnSerProLysLeu 180
      486      CAGATGTAGACTAGCTGCTGATGCTGGTATCAACAAACAGGCGCATCTCCAAACTA 545
QY      181      LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
      546      CTGATTTAATCTGACATCCACCGGACACCTGAGTCCCTGATCCCTTCACAGCGAGTGA 605
QY      201      SerGlyProAspPheThrPheThrIleSerSerValGlnIleAsnIleValTyr 220
      606      TCTGGGACAGATTATCTTCACATCAGACAGTGTGAGGCTGAAGACCTGGCACTTCAT 665
QY      221      PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyTyrHisLeuGlnIle 240
      666      TACTGTACAGCAACTATATAGAGTCCGACACGTTCCGAGGCGGAGCCAGAGCTGGAATA 725
QY      241      Lys 241
      726      AAA 728
Db

```

```

RESULT 12
US-08-461-184-7
; Sequence 7, Application US/08461184
; Patent No. 5631158
; GENERAL INFORMATION:
; APPLICANT: DORAI, HAIMANTI
; APPLICANT: OPERMANN, HERMANN
; TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN
; TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC
; STREET: 45 SOUTH STREET
; CITY: HOPKINTON
; STATE: MA
; COUNTRY: USA
; ZIP: 07148
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,184
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/143,498

```

```

; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: KELLEY, ROBIN D
; REGISTRATION NUMBER: 34,637
; REFERENCE/DOCKET NUMBER: CRP093
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/248-7000
; TELEFAX: 617/248-7100
; INFORMATION FOR SRO ID NO.: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 909 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 3..752
; OTHER INFORMATION: /product= "7A1F6 sfv"
US-08-461-184-7

Alignment Scores:
Pred. No.: 1.84e-85 Length: 909
Score: 981.50 Matches: 185
Percent Similarity: 85.89% Conservative: 22
Best Local Similarity: 76.76% Mismatches: 33
Query Match: 41.86% Indels: 1
DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-461-184-7 (1-909)
QY      1      GlnIleGlnLeuValGlnSerGlyProGluLeuLysPheProGlyGluThrValLysIle 20
      9      GAATCCAAATGGTGTGACGTCTGACCTGACCTGACAGAACCTGGAGAGACATGCAATC 68
QY      21      SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40
      69      TCTGCAAGGCTTCTGGGTATACCTTCACAAACTATGATGAATGAACCTGGTGAAGCAGGCT 128
QY      41      ProGlyGlnGlyLeuLeuTyrPheMetGlyTyrPheLeuThrSerThrGlyGluSerThrPhe 60
      129      CCAGGAAGGCTTTAAAGTGAATGAGGCTGATTAACCAACCACTGGAGAGCCAAATAT 188
QY      61      AlaAspAspPheLeuGlyValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
      189      GCTGAAGAGTTCAAGGAGCGTTTGCTTCTTGGAAACCTTCGCCACAGCTGCTAT 248
QY      81      LeuGlnIleAsnAsnLeuLeuSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
      249      TTGCAGATCAACAACCTCAAAATAGAGACACGGCTACATATTTCTGTGGAAGCAATT 308
QY      101      ValTyrHisGlyTyrValProTyrTyrPheGlyGlnGlyThrThrValThrValSerSerGly 120
      309      ATTACCTACGCGGGTTTGCTACCTGGGCGCAAGGAGTCTGCTGCTGCTGCA--- 365
QY      121      GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGln 140
      366      TCGAGCTCCTCCGATCTTCATCTAGCGGTTCCAGCTCGACGATATCGTCATGACCCAG 425
QY      141      SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAlaSer 160
      426      TCTCTTAATTCATGCTCCACGCTCAGTGGAGACAGGCTCAGCATCTCTGCAAGGCCAGT 485
QY      161      GlnAspValTyrAsnAlaValAlaTyrPheGlnGlnLysProGlyGlnSerProLysLeu 180
      486      CAGATGTAGACTAGCTGCTGATGCTGATCAACAAACAGGCGCATCTCTTAAACTA 545
QY      181      LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
      546      CTGATTTAATCTGACATCCACCGGACACCTGAGTCCGATCGCTTCAAGGAGCTGA 605
QY      201      SerGlyProAspPheThrPheThrIleSerSerValGlnIleGlnAspLeuAlaValTyr 220

```

Db 606 TCTGGACAGATATTAATCTCAACCATGACAGTGTGACAGCTGAAGACCTGGCACTTCAT 665  
Qy 221 PheCyGlnGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrValLeuGlnIle 240  
Db 666 TACTGTACGACAACTATTATAGTGCCTGACACGTTTCGGAGGGGACCAAGCTGGAGATA 725  
Qy 241 Lys 241  
Db 726 AAA 728

RESULT 13  
US-08-463-675-7  
Sequence 7, Application US/08463675  
Patent No. 5658763  
GENERAL INFORMATION:  
APPLICANT: DORAI, HAIMANTI  
APPLICANT: OPPERMANN, HERMANN  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN  
TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 07148

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/463,675  
FILING DATE: 05-JUN-1995  
CLASSIFICATION: 435  
PRIOR APPLICATION DATA:  
APPLICATION NUMBER: 08/143,498  
FILING DATE: 25-OCT-1993  
ATTORNEY/AGENT INFORMATION:  
NAME: KELLEY, ROBIN D  
REGISTRATION NUMBER: 34,637  
REFERENCE/DOCKET NUMBER: CRP093  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: 617/248-7000  
TELEFAX: 617/248-7100  
INFORMATION FOR SEQ ID NO: 7:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 909 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
FEATURE:  
NAME/KEY: CDS  
LOCATION: 3..752  
OTHER INFORMATION: /product= "741P8 sFv."

US-08-463-675-7

Alignment Scores:  
Pred. No.: 1.84e-85 Length: 909  
Score: 981.50 Matches: 185  
Percent Similarity: 85.8% Conservative: 22  
Best Local Similarity: 76.7% Mismatches: 33  
Query Match: 41.86% Indels: 1  
DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-463-675-7 (1-909)

Qy 1 GlnIleGlnLeuValGlnSerGlyProGlnLeuValSerProGlyGlnThrValIle 20  
Db 9 GAGATCCAAATGTGTGACGTGACCTGAGCTGAAGAGCCCTGGAGAGACAGTCAAGATC 68

Qy 21 SerCyValAlaSerGlyThrProPheThrAsnTyrGlyMetAsnTyrValIleGlnAla 40  
Db 69 TCTTGCAAGGCTTCTGGGTATACCTTCACAAACTATGAAATGAACTGGGTGACAGGCT 128  
Qy 41 ProGlyGlnGlyLeuValSerPheGlyTyrIleAsnThrSerThrGlySerThrPhe 60  
Db 129 CCAGGAAAGGGGTTTAAAGTGTATGGCTGTGATTAACCAACACTGGAGACCAACATAT 188  
Qy 61 AlaAspAspPheValSerGlyPheAspPheSerLeuGlnThrSerAlaAsnThrAlaTyr 80  
Db 189 GCTGAAGAGTTCAGAGGACGGATTGCTTCTTTGAAACCTTCTCCAGACAGCTGCTAT 248  
Qy 81 LeuGlnIleAsnAsnLeuValSerGlyAspMetAlaThrTyrPheCyAlaArgTyrGln 100  
Db 249 TTGCAGATCAACAACTCCAAATGAGAGACAGGCTCATATTTCTGTGAAGCAATTT 308  
Qy 101 ValTyrHisGlyTyrValProTyrTyrPheGlnGlyThrValThrValSerSerGly 120  
Db 309 ATTACTACGCGCGGTTTGCTACTGGGGCCAGGACTGTGTCACTGTCTGCA--- 365  
Qy 121 GlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerAspIleGlnLeuThrGln 140  
Db 366 TCGAGCTCTCCGATCTTCATCTAGCGGTTCCAGCTCGAGCGATATGTCATACCCAG 425  
Qy 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCyValAlaSer 160  
Db 426 TCTCTTAATTCATGCTCACTCATGCTGAGGAGACAGGCTCAGCATCTCTGCAAGGCCAGT 485  
Qy 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180  
Db 486 CAGAGTGTGAGTACTGCTGTAGCTGATTCACAAACCAAGGAGCAATCTCTAACTA 545  
Qy 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
Db 546 CTGATTACTGACATCCACCCGCGACACCTGAGTCTCTGATCTCTTACACGAGGTGA 605  
Qy 201 SerGlyProAspPheThrPheThrIleSerSerValGlnIleGlnAspLeuAlaValTyr 220  
Db 606 TCTGGACAGATTTATCTCACTCACTGACAGTGTGAGGCTGAAGACCTGGCACTTCAT 665  
Qy 221 PheCyGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrValLeuGlnIle 240  
Db 666 TACTGTACGACAACTATTATAGTGCCTGACACGTTTCGGAGGGGACCAAGCTGGAGATA 725  
Qy 241 Lys 241  
Db 726 AAA 728

RESULT 14  
US-08-464-589-7  
Sequence 7, Application US/08464589  
Patent No. 5733782  
GENERAL INFORMATION:  
APPLICANT: DORAI, HAIMANTI  
APPLICANT: OPPERMANN, HERMANN  
TITLE OF INVENTION: METHODS AND COMPOSITIONS FOR HIGH PROTEIN  
TITLE OF INVENTION: PRODUCTION FROM NON-NATIVE DNA  
NUMBER OF SEQUENCES: 10  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: PATENT ADMINISTRATOR, CREATIVE BIOMOLECULES, INC  
STREET: 45 SOUTH STREET  
CITY: HOPKINTON  
STATE: MA  
COUNTRY: USA  
ZIP: 07148

COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: Patentin Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/464,589  
FILING DATE: 05-JUN-1995

CLASSIFICATION: 435  
 PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US 08/143,498  
 FILING DATE: 25-OCT-1993  
 ATTORNEY/AGENT INFORMATION:  
 NAME: KELLEY, ROBIN D  
 REGISTRATION NUMBER: 34,637  
 REFERENCE/DOCKET NUMBER: CRP093  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617/248-7000  
 TELEFAX: 617/248-7100  
 INFORMATION FOR SEQ ID NO: 7:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 909 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA  
 FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 3..752  
 OTHER INFORMATION: /product= "741F8 sFv"  
 US-08-464-589-7

Alignment Scores:  
 Pred. No.: 1,84e-85 Length: 909  
 Score: 981.50 Matches: 185  
 Percent Similarity: 85.89% Conservative: 22  
 Best Local Similarity: 76.76% Mismatches: 33  
 Query Match: 41.86% Indels: 1  
 DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-464-589-7 (1-909)

QY 1 GlnlleglnleuValGlnserGlyProGlnleuLysPProGlyGluThrValLyslle 20  
 Db 9 GAGATCCAAATGGGCTGACCTGACCTGAGCTGAAGAGCCTGAGAGACGTCMAAGTC 68  
 QY 21 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTPvalLysGlnAla 40  
 Db 69 TCCGCAAGGCTTGGGTATACCTTCACAACTATGATGAACTGGGAGAACAGGCT 128  
 QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrPheAsnThrSerThrGlyGluSerThrPhe 60  
 Db 129 CCAGGAAAGGCTTAAAGTGATGGCTGGATTAACCAACAACATGAGAGCCAACTAT 188  
 QY 61 AlaSpaSerPheLysGlyTyrGlyPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80  
 Db 189 GCTGAAGAGTTCAAGGACGGCTTGCCTTCCTTGGAAACCTCTGCCAGCACTGCCTAT 248  
 QY 81 LeuGlnIleAsnAsnLeuLysSerGlyAspPheAlaThrTyrPheCysAlaArgTyrGlu 100  
 Db 249 TTGCAGATCAACAACCTCAAAAATGAGACACGGCTACATATTTCTGGAAGGCAATT 308  
 QY 101 ValTyrHisGlyTyrValProTyrTyrPylGlnGlyThrValThrValLysSerGly 120  
 Db 309 ATTACCTACGGCGGCTTGGTAACCTGGGGCCAGGAGCTCGTCACTGCTCTGCA--- 365  
 QY 121 GlyGlyLysSerGlyGlyGlyLysSerGlyGlyLysSerAspIleGlnLeuThrGln 140  
 Db 366 TCGAGCTCTCCGGATCTTCATCTAGCGGTTCCAGCTCGAGCATATCGCATACCCAG 425  
 QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160  
 Db 426 TCTCTAAATTCATGCTCACTGAGGAGACAGGGCTCAGCATCTCTCAAGGCGCAGT 485  
 QY 161 GlnAspValLysPheAsnAlaValAlaTyrPylGlnGlnLysPProGlyGlnSerProLysLeu 180  
 Db 486 CAGAGTGTGAGTACTGCTAGCTGTATCAACAAACCAAGGCAATCTCTAAACTA 545  
 QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200  
 Db 546 CTGATTTACGACATCCACCGGACACATGAGTCCCTGATCGCTTCACAGGACAGTGA 605

QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220  
 Db 606 TCTGGACAGATTATCTCATCCATCAAGAGTGTGAGGCTGAAGACCTGGACCTTCAT 665  
 QY 221 PheCysGlnGlnHisPheArgTyrProPheThrPheGlySerGlyThrLysLeuGlnIle 240  
 Db 666 TACTGTAGCACTATTATAGTGCCTTACACGTTCCGAGGGGGACCAAGCTGGAGATA 725  
 QY 241 Lys 241  
 Db 726 AAA 728

# RESULT 15

US-08-461-838-1  
 Sequence 1, Application US/08461838  
 Patent No. 5753204  
 GENERAL INFORMATION:  
 APPLICANT: Huston, James S.  
 APPLICANT: Oppermann, Hermann  
 APPLICANT: Houston, L. L.  
 TITLE OF INVENTION: Ring, David B.  
 TITLE OF INVENTION: Biosynthetic Binding Proteins For  
 NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESSEE: Testa, Hurwitz & Thibault/Patent Department  
 STREET: Exchange Place, 53 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: Floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent In Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/08/461,838

FILING DATE:  
 CLASSIFICATION: 424  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kelley, Robin D.  
 REGISTRATION NUMBER: 34,637  
 REFERENCE/DOCKET NUMBER: 2054/22  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-248-7477

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 909 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 3..752  
 OTHER INFORMATION: /product= "741F8 sFv" C-terminal  
 OTHER INFORMATION: Gly4-Cys"  
 US-08-461-838-1

Alignment Scores:  
 Pred. No.: 1,84e-85 Length: 909  
 Score: 981.50 Matches: 185  
 Percent Similarity: 85.89% Conservative: 22  
 Best Local Similarity: 76.76% Mismatches: 33  
 Query Match: 41.86% Indels: 1  
 DB: 1 Gaps: 1

US-09-596-774-7 (1-443) x US-08-461-838-1 (1-909)

QY 1 GlnlleglnleuValGlnserGlyProGlnleuLysPProGlyGluThrValLyslle 20

```
Db      ::::|||||
9      GAGATCCAAATGGTGCAGCTGAGCCTGAGAAAGCCTGAGAGACAGTCAAGATC 68
QY      21 SerCysIysAlaSerGlyTyrProPheThrAenTYrGlyMetAenTrpValIysGlnAla 40
Db      69 TCCTGCAAGGCTTCGGGTATACCTTCACAAACTATGATGATGACTGGGTGAAGCAGGCT 128
QY      41 ProGlyGlnGlyLeuIlyTrpMetGlyTyrPleAenThrSerThrGlyIlySerThrPhe 60
Db      129 CCAGGAAAGGTTTAAAGTGAATGGGCTGGATTAACCAACCACTGAGAGCCACATAT 188
QY      61 AlaAspAspPheIlyGlyArgPheAspPheSerLeuGluThrSerAlaAenThrAlaTyr 80
Db      189 GCTGAAGAGTTCAAGAGCAGGTTTGCTCTCTTGGAAACCTGCGCAGCAGCTGCTAT 248
QY      81 LeuGlnIleAenAenLeuIlySerGlyAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db      249 TTGCAGATCAACMACCTCAAAATGAGGACACGGCTACATATTTCTGGAAAGCCAAATT 308
QY      101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
Db      309 ATTACCTACGGCGGTTGGTCTACTGGGCCAAAGGAACTTGCTCATCTGTCTGCA--- 365
QY      121 GlyGlyIlySerGlyGlyGlyIlySerGlyIlyGlySerAspIleGlnLeuThrGln 140
Db      366 TCGAGCTCCTCGGATCTTCATCTAGCGGTTCCAGCTCAGCGATATCGTATGACCCAG 425
QY      141 SerHisIlyPheLeuSerThrSerValGlyAspArgValSerIleThrCysIlyAlaSer 160
Db      426 TCTCCTTAATTCATGTCCACGTCAGTGGGAGACAGGTCAGCATCTCTGCAAGGCCAGT 485
QY      161 GlnAspValTyrAsnAlaValAlaTrpIlyGlnGlnIlyProGlyGlnSerProIlySer 180
Db      486 CAGATGTGAGTACTGCTGTAGCTGTGATCAACAAAACAGGCAATCTCTTAACTA 545
QY      181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db      546 CTGATTTACTGGAACATCCACCGGACACTGGAGTCCCTGATCGCTTCACAGGCAGTGSA 605
QY      201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
Db      606 TCTGGGACAGATTATCTCTCACCATGACAGTGTGCAAGGCTGAAAGACCTGGCACTTCAT 665
QY      221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrIlyLeuGluIle 240
Db      666 TACTGTCAACAACATTATAGAGTCCGTACACGTTTGGAGGGGGGACCAAGCTGAGATA 725
QY      241 Lys 241
Db      726 AAA 728
```

Search completed: April 20, 2005, 16:41:44  
Job time : 239.44 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 20, 2005, 10:19:55 ; Search time 758.098 Seconds

(without alignments)  
3545.124 Million cell updates/sec

Title: US-09-596-774-7

Perfect score: 2345

Sequence: 1 QIQLVQSGPELKKPEFTVKI.....LSTATKDYDALHMQTLAPR 443

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 5622541 seqs, 303355566 residues

Total number of hits satisfying chosen parameters: 11245082

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:  
-MODEL=frame+g2n.model -DEV=xlh  
-Q=/cgn2\_1/USPTO.spool\_h/US09596774/runat\_20042005\_075444\_12560/app.query.fasta\_1.1230  
-DB=Published Applications NA -QPM=fastap -SUFFIX=apr20.rnpb -MINMATCH=0.1  
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62  
-TRANS=human40.cdd -LIST=45 -DOCLIGN=200 -THR SCORE=pct -THR MAX=100  
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0  
-MAJLEN=200000000 -USER=US09596774@CGN\_1.698@runat\_20042005\_075444\_12560  
-NCPU=6 -ICPU=3 -NO MAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSELOCK=100  
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5  
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published Applications NA:\*

- 1: /cgn2\_6/ptodata/1/pubpna/US07\_PUBCOMB.seq:\*
- 2: /cgn2\_6/ptodata/1/pubpna/PCT\_NEW\_PUB.seq:\*
- 3: /cgn2\_6/ptodata/1/pubpna/US06\_NEW\_PUB.seq:\*
- 4: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 5: /cgn2\_6/ptodata/1/pubpna/US07\_NEW\_PUB.seq:\*
- 6: /cgn2\_6/ptodata/1/pubpna/PCTUS\_PUBCOMB.seq:\*
- 7: /cgn2\_6/ptodata/1/pubpna/US08\_NEW\_PUB.seq:\*
- 8: /cgn2\_6/ptodata/1/pubpna/US08\_PUBCOMB.seq:\*
- 9: /cgn2\_6/ptodata/1/pubpna/US09\_PUBCOMB.seq:\*
- 10: /cgn2\_6/ptodata/1/pubpna/US09B\_PUBCOMB.seq:\*
- 11: /cgn2\_6/ptodata/1/pubpna/US09C\_PUBCOMB.seq:\*
- 12: /cgn2\_6/ptodata/1/pubpna/US09\_NEW\_PUB.seq:\*
- 13: /cgn2\_6/ptodata/1/pubpna/US10A\_PUBCOMB.seq:\*
- 14: /cgn2\_6/ptodata/1/pubpna/US10B\_PUBCOMB.seq:\*
- 15: /cgn2\_6/ptodata/1/pubpna/US10C\_PUBCOMB.seq:\*
- 16: /cgn2\_6/ptodata/1/pubpna/US10D\_PUBCOMB.seq:\*
- 17: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 18: /cgn2\_6/ptodata/1/pubpna/US10F\_PUBCOMB.seq:\*
- 19: /cgn2\_6/ptodata/1/pubpna/US10F\_NEW\_PUB.seq:\*
- 20: /cgn2\_6/ptodata/1/pubpna/US11\_NEW\_PUB.seq:\*
- 21: /cgn2\_6/ptodata/1/pubpna/US60\_NEW\_PUB.seq:\*
- 22: /cgn2\_6/ptodata/1/pubpna/US60\_PUBCOMB.seq:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1249	53.3	1920	US-10-120-198B-1	Sequence 1, Appl1
2	1249	53.3	6834	US-10-120-198B-5	Sequence 5, Appl1
3	1015.5	43.3	7654	US-10-006-773-1	Sequence 1, Appl1
4	1015.5	43.3	7654	US-10-006-773A-1	Sequence 1, Appl1
5	994	42.4	771	US-10-013-173-38	Sequence 38, Appl1
6	994	42.4	771	US-10-150-762-38	Sequence 38, Appl1
7	994	42.4	771	US-10-244-821-38	Sequence 38, Appl1
8	981.5	41.9	909	US-09-887-853-1	Sequence 1, Appl1
9	969.5	41.3	909	US-09-887-854-1	Sequence 1, Appl1
10	930	39.7	1350	US-08-812-393A-1	Sequence 1, Appl1
11	928	39.6	1350	US-09-774-681-1	Sequence 1, Appl1
12	917.5	39.1	1515	US-10-239-656-78	Sequence 78, Appl1
13	913	38.9	816	US-09-766-543-9	Sequence 9, Appl1
14	903.5	38.5	723	US-10-127-890-90	Sequence 90, Appl1
15	903.5	38.5	723	US-10-717-243-90	Sequence 90, Appl1
16	897	38.3	840	US-09-766-543-11	Sequence 11, Appl1
17	889.5	37.9	1626	US-10-071-485-84	Sequence 84, Appl1
18	888.5	37.9	804	US-10-071-485-1	Sequence 1, Appl1
19	888.5	37.9	2133	US-10-071-485-89	Sequence 89, Appl1
20	886.5	37.8	1509	US-10-239-656-74	Sequence 74, Appl1
21	875.5	37.3	10511	US-10-059-261-109	Sequence 109, App
22	875.5	37.3	10511	US-10-627-649-109	Sequence 109, App
23	872.5	37.2	729	US-10-879-994-9	Sequence 9, Appl1
24	872.5	37.2	729	US-10-610-452-9	Sequence 9, Appl1
25	860.5	36.7	761	US-10-354-246-4	Sequence 4, Appl1
26	859	36.6	1869	US-10-378-832A-1	Sequence 1, Appl1
27	854.5	36.4	1509	US-10-239-656-76	Sequence 76, Appl1
28	854	36.4	777	US-10-096-246-5	Sequence 5, Appl1
29	854	36.4	777	US-10-096-246-7	Sequence 7, Appl1
30	849.5	36.2	1497	US-10-239-656-72	Sequence 72, Appl1
31	848	36.2	702	US-10-071-485-92	Sequence 92, Appl1
32	846	36.1	777	US-10-096-246-6	Sequence 6, Appl1
33	841.5	35.9	756	US-09-984-186-17	Sequence 17, Appl1
34	841.5	35.9	756	US-10-237-667-17	Sequence 17, Appl1
35	841.5	35.9	756	US-10-237-708-17	Sequence 17, Appl1
36	841.5	35.9	756	US-10-237-866-17	Sequence 17, Appl1
37	841.5	35.9	756	US-10-237-871-17	Sequence 17, Appl1
38	841.5	35.9	756	US-10-237-824-17	Sequence 17, Appl1
39	841.5	35.9	756	US-10-702-536-17	Sequence 17, Appl1
40	841.5	35.9	756	US-10-702-636-17	Sequence 17, Appl1
41	830.5	35.4	739	US-09-887-853-5	Sequence 5, Appl1
42	830.5	35.4	739	US-10-683-547-5	Sequence 5, Appl1
43	830.5	35.4	774	US-10-096-246-3	Sequence 3, Appl1
44	827.5	35.3	723	US-10-335-290-169	Sequence 169, App
45	826	35.2	921	US-10-363-349-12	Sequence 12, Appl1

## ALIGNMENTS

RESULT 1  
US-10-120-198B-1  
Sequence 1, Application US/10120198B  
Publication No. US20030215427A1  
GENERAL INFORMATION: Jensen, Michael  
APPLICANT: Jensen, Michael  
TITLE OF INVENTION: CEF-SPECIFIC REDIRECTED IMMUNE CELLS  
FILE REFERENCE: 1954-337  
CURRENT APPLICATION NUMBER: US/10/120,198B  
CURRENT FILING DATE: 2002-04-11  
PRIOR APPLICATION NUMBER: 60/282,859  
PRIOR FILING DATE: 2001-04-11  
NUMBER OF SEQ ID NOS: 11  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1920  
TYPE: DNA  
ORGANISM: artificial sequence  
FEATURE:  
OTHER INFORMATION: mouse-human chimera

FEATURE:  
NAME/KEY: CDS  
LOCATION: (14) (1906)  
OTHER INFORMATION: scfvfc construct  
US-10-120-198B-1

Alignment Scores:  
Pred. No.: 1.19e-120 Length: 1920  
Score: 1249.00 Matches: 276  
Percent Similarity: 54.60% Conservative: 62  
Best Local Similarity: 44.59% Mismatches: 96  
Query Match: 53.26% Indels: 186  
DB: 17 Gaps: 12

US-09-596-774-7 (1-443) x US-10-120-198B-1 (1-1920)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGlyGluThrValLysIle 20
DB CAGGTCCCAACTGCAGCAGCCTGGGGCTGACTGGTGAAGCCTGGGGCTTCAGTGAAGCTG 139
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetLeuThrValLysGlnAla 40
DB TCCTGCAAGGCTTCTGCTAACCCTTACCGGCTACTGGATGCACTGGTGAAACGAGG 199
QY 41 ProGlyGlnGlyLeuLysTyrPheGlyTyrPheLeuThrSerThrGlyGluSerThrPhe 60
DB CCTGGACATGGCTTGAAGTGAATGAGATTAATCCTACCAACGGTCGTAACACTAC 259
QY 61 AlaAspAspPheLysGlyLysArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB AATAAGAGTTCACAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 319
QY 81 LeuGlnIleAsnLeuLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTyrGlu 100
DB ATGCACTCAGCAGCGGCTTACATCTGAGACTCTGCACTATTTCTGTGCAAGA----- 373
QY 101 ValTyrHisGly-----TyrValProTyrTyrGlyGlnGlyThrThrValThrVal 117
DB GATTACTACGGTACTAGCTACCACTTGTGACTACTGGGGCCAAAGCAGCAGCTCAGAGTC 433
QY 118 SerSerGlyGlyGlySerGlyGlyGlyGlyGlySerGlyGlyGlySerAspIleGln 137
DB TCCCTCAGAGGTGGCGGTAGTGAAGTGGCGGATCCGGTGGCGGAGTGAAGCATCTCAG 493
QY 138 LeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCys 157
DB ATGACACATCTTCACTCTCTTTCTGTATCTCTAGAGACAGAGTCAACATTACTTGC 553
QY 158 LysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSer 177
DB AAGGCTAATGAAGCATAAATATCGGTAGCTGTATCAGCAGACACCGAATAATCT 613
QY 178 ProLysLeuLeuLysTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThr 197
DB CTAGGCTCTTAAATATGTGTGCAACCAATTGGTAAGTGGGGTTCCTTCAAGATTGAGT 673
QY 198 GlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnIleGluAspLeu 217
DB GCGAGTGAAGTGAAGGATTAACCTCAACCTTACCAAGTCTTCAAGGCTGAAGATTTT 733
QY 218 AlaValTyrPheCysGlnGlnHisAspArgThrProPheThrPheGlySerGlyThrLys 237
DB GCTACTATTATCTGTCAACATATTTGAGTACTCATTCAAGTTCCGCTCGGGGACAGAG 793
QY 238 LeuGluLys----- 241
DB CTGAGATCAAAAGTAGAACCAAACTTCTGACAAAACTCACACATGCCCAACCGTGCCA 853
QY 241 ----- 241
DB GCACTGAAGTCTTGGGGGAGCCTCAGTCTTCTCTTCCCCCAAAACCAAGAGACACC 913
QY 242 -----AlaLeuGluLysSer----- 246

```

```

DB 914 CTGATGATCTCCGGAGCCCTGAGGTCAATGCGTGTGTGAGCGTGAACCAAGAAC 973
QY 246 ----- 246
DB 974 CTTGAGGTCAAGTTCACTGTGATCGTGAAGCGGTGGAGGTGATTAATGCAAGACAAAG 1033
QY 247 -----AsnSerValMetTyrPheSerSerValProValIleGln 260
DB 1034 CCGCGGAGAGAGATCAACACAGACGTAACGTTGTGTGCTCAGCTTCTCAGCTTCGAC 1093
QY 261 -----LysValAsnSerThrThrThrLysPhe 269
DB 1094 CAGGACTGCTGAATGCAAGAGTACACTGCAAGGTCTCCAAACAAAGCCCTCCAGCC 1153
QY 269 oValLeuArgThrProSerProValHisProThrGlyThrSerGlnProGlnArgProGln 289
DB 1154 CCCATCGAAGAAACCATCTCCA---AAGCCAAGGGAGCCCGAAGAACACAGATGTAC 1210
QY 289 uAspCys----- 291
DB 1211 ACCCTGCCACCATCAGAGATGAGTGAACCAAGAACAGGTCACTGACCTGCTGCTGTC 1270
QY 292 -----ArgProArgGlySerValLysGly 299
DB 1271 AAAGGCTTATCCAGGACGATCGCGGTGAGTGGAGAGCATGGGACCGGAGAAC 1330
QY 299 YThrGlyLeuAspPheLeuGlyLysProLysLeu---CysTyrLeuLeuAspGlyIleLe 318
DB 1331 AACT-----ACAAAGACACGCTCCGCTGTGATCTCC---GACGGCTCTT 1374
QY 318 uPheIleTyr----- 321
DB 1375 CTTCCTCTACAGCAAGCTACCGTGAACAGAGACAGAGTGGACAGAGGAGACGTCTTCTC 1434
QY 321 ----- 321
DB 1435 ATGCTCGTATGATGATGAGGCTGTGACAAACATACAGCAGAAAGCCTGCTGCTGTC 1494
QY 322 -----GlyValIleIle----- 325
DB 1495 TCCCGGAAAATGCGCTGATGCTGCGGGGCGTGGCGGCTCTCTCTTTCATTGG 1554
QY 326 -ThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThrAlaAlaAsnLeuGly 345
DB 1555 GCTAGGCACTCTTTCAGAGTGAAGTTCAGCAGAGCGAGACGCGCCCGCTACACAGA 1614
QY 345 nAspProAsnGlnLeuTyrAsnGluLeuAsnLeuGlyArgTyrGluTyrAspValLe 365
DB 1615 GGGCAGAAACAGCTCTATTAACGAGCTCAATCTAGAGCAGAAAGAGAGATGATTTT 1674
QY 365 uGluLysLysArgAlaArgAspProGluMetGlyLysGlnGlnArgTyrArgAspArg 385
DB 1675 GGCACAAAGACGCGGGCGGAGCCCTGAGATGGGGGAAAG---CCGAGAAAGAAAGAACCC 1731
QY 385 oGlnGluGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGluAlaTyrSerGluIle 405
DB 1732 TCGAGAGGCTCTTACAAAGAACTGCAAGAAAGTAAAGTGAAGGAGGCTTACAGTGAAT 1791
QY 405 eGlyThrLysGlyLysArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSe 425
DB 1792 TGGAGTGAAGAGGAGCGCGCGGAGGGGCAAGGGGACGATGCGCTTTCACAGGCTCGAG 1851
QY 425 rThrAlaThrLysAspThrTyrAspAlaLeuHisMetGlnThrLeuAlaProArg 443
DB 1852 TACAGCCACCAAGAGACCTTACGACGCCCTTCAATGACGGCCCTGCCCTCCGCGC 1906

```

## RESULT 2

US-10-120-198B-5  
Sequence 5, Application US/10120198B  
Publication No. US20030215427A1  
GENERAL INFORMATION:  
APPLICANT: Jensen, Michael



; TITLE OF INVENTION: CEF7-SPECIFIC REDIRECTED IMMUNE CELLS  
 ; FILE REFERENCE: 1954-337  
 ; CURRENT APPLICATION NUMBER: US/10/120,198B  
 ; PRIOR FILING DATE: 2002-04-11  
 ; PRIOR APPLICATION NUMBER: 60/282,859  
 ; NUMBER OF SEQ ID NOS: 11  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 5  
 ; LENGTH: 6834  
 ; TYPE: DNA  
 ; ORGANISM: artificial sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: plasmid construct  
 ; US-10-198B-5

Alignment Scores:

Prod. No.:	6.65e-120	Length:	6834
Score:	1249.00	Matches:	276
Percent Similarity:	54.60%	Conservative:	62
Best Local Similarity:	44.59%	Mismatches:	96
Query Match:	53.26%	Indels:	186
DB:	17	Gaps:	12

US-09-596-774-7 (1-443) x US-10-120-198B-5 (1-6834)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuValSerProGlyGluThrValLysIle 20  
 DB 1878 CAGGTCACACGACGACGCTGGGCTGACCTGGAAGAGCTGGGGCTTCAGTGAAGCTG 1937  
 QY 21 SerCysValAspAlaSerGlyTyrProPheThrAsnTyrGlyMetLeuAspTyrValLysGlnAla 40  
 DB 1938 TCCTGCAAGGCTTCTGGCTACACCTTCAACCGGCTACCTGATGCTGAGGTGAAGCAGAG 1997  
 QY 41 ProGlyGlnGlyLeuLeuTyrPheGlyTyrPheLeuThrSerThrGlyGluSerThrPhe 60  
 DB 1998 CCTGGACATGCGCTTGGATGATGAGATGAGATTAATCTTACGCAAGCTGCTACTAATCTAC 2057  
 QY 61 AlaAspAspPheLeuValArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80  
 DB 2058 AATGAGAGCTTCAGAGCAAGGCGCACCTGACTGAGCAAAATCTCCACCAACGCTTC 2117  
 QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrPheGlu 100  
 DB 2118 ATGCACTACACGCGCTGACATCTGAGACTCTGCAAGCTAATTTCTGCAAGA----- 2171  
 QY 101 ValTyrHisGly-----TyrValProTyrTyrPheGlyGlnGlyTyrThrValThrVal 117  
 DB 2172 GATTACTACGGTACTAGCTACCAACTTGTACTGAGGCGCAAGGCAACCACTCTCAGCTC 2231  
 QY 118 SerSerGlyGlyGlyGlySerGlyGlyGlySerGlyGlyGlySerGlyGlySerAspIleGln 137  
 DB 2232 TCCTCAGAGAGGTGGGAGTGGAGAGTGGCGGATCCGGTGGCGGAGTGGAGCACTTCAG 2291  
 QY 138 LeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCys 157  
 DB 2292 ATGACACATCTTCATCTCTCTTCTGTATCTCTGAGAGACAGAGTCAACCTTACTTGC 2351  
 QY 158 LysAlaSerGlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSer 177  
 DB 2352 AAGCTAATGAAACATTAATTAATCGTGTGATGCTGATCAGACAGACCAAGAAATCT 2411  
 QY 178 ProLysLeuLeuLeuTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThr 197  
 DB 2412 CCTAGGCTCTTAATATCTGCTGCAACCAATTTGTACTGGGGTCTTCAAGATTCGT 2471  
 QY 198 GlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeu 217  
 DB 2472 GCGAGTGAATCTGAAAGATTACCTCTCACCATTACCAAGCTTCAAGCTGAAGATTCTT 2531  
 QY 218 AlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrLys 237  
 DB 2532 GCTACTTATTAATGCAACAAATATTGAGTACTCATTCACGTTGGGCTGGGAGCAAGAG 2591

QY 238 LeuGlnIleLys----- 241  
 DB 2592 CTCGAGATCAAGATGAAGAACCCAAATCTTGTGACAAAACACTACACATGCCACCGGCCCA 2651  
 QY 241 ----- 241  
 DB 2652 GCACCTGAATCTGGGGGACCGTCACTTCTCTTCCCCCAAAACCAAGAGACACC 2711  
 QY 242 -----AlaLeuGlnIleSer----- 246  
 DB 2712 CTCATGATCTCCGGAGCCCTGAGGTCAATGCTGTGTGAGCAGTGAAGCAGAGAGAC 2771  
 QY 246 ----- 246  
 DB 2772 CTTGAGTCAAGTTCACTGCTGATGACGCGCTGAGGTGCATATGCCAAGCAAG 2831  
 QY 247 -----AsnSerValMetTyrPheSerSerValValProValLeuGln 260  
 DB 2832 CCGGCGAGAGAGAGTACACAGACAGCGTGTGCTGCTCCTCAGCGTCTGAC 2891  
 QY 261 -----LysValAsnSerThrThrThrLysPhe 269  
 DB 2892 CAGGACTGGCTGAATGGCAAGAGTACAAAGTCTCCAAACAAAGCCCTCCACACC 2951  
 QY 269 ovalLeuArgThrProSerProValHisProThrGlyTyrSerGlnProGlnArgProGln 289  
 DB 2952 CCATTCGAGAAACATCTTCCA--AACCAAGGCGAGCCCGAGAAACACAGGTGAC 3008  
 QY 289 uAspCys----- 291  
 DB 3009 ACCCTGCACATCAGACGATGAGGTGACCAAGAACAGTCAAGCTGACCTGCTGTC 3068  
 QY 292 -----ArgProArgGlySerValLysGln 299  
 DB 3069 AAAGCTTCTATCCAGAGCATGCGCTGAGTGGAGAGCAATGGGACACCGGAGAAC 3128  
 QY 299 YThrGlyLeuAspPheLeuGluAspProLysLeu-----CysTyrLeuLeuAspGlyIleLe 318  
 DB 3129 AACCT-----ACAAAGCAGCGCTCCGCTGCTGAGCTCC-GACGCTCTCT 3172  
 QY 318 uPheIleTyr----- 321  
 DB 3173 CTTCTCTACAGCAAGCTCACCGTGAACAGAGCAGGTGCGAGGAGAAAGTCTTCTC 3232  
 QY 321 ----- 321  
 DB 3233 ATGCTCCGTATGATGAGGCTCTGACAAACCACTACACGAGAAAGCTCTCTGTC 3292  
 QY 322 -----GlyValIleIle----- 325  
 DB 3293 TCCCGGAGAAATGCGCTGATTTGCTGGGGGGGCTCGCCCGCTCTCTTTCATTGG 3352  
 QY 326 -ThrAlaLeuTyrLeuArgAlaLysPheSerArgSerAlaGluThrAlaAlaAsnLeuGln 345  
 DB 3353 GCTAGGCACTTCTTCAAGATGAAGTTTCAAGAGAGGCGAGACCGCCCGGCTACAGCA 3412  
 QY 345 nAspProAsnGlnLeuTyrAsnGlnLeuAsnLeuGlyArgArgGlnGluTyrAspValLe 365  
 DB 3413 GGGCGAGAACCGCTCTAATACGAGCTCAATCTGAGCGAAGAGAGGAGTGCAGTGTCTT 3472  
 QY 365 uGluLysLeuArgAlaArgAspProGluMetGlyGlyLysGlnGlnArgArgAspAsp 385  
 DB 3473 GAGCAAGAGAGCGCGCGGACCTGAGATGGGGGAGAG--CCGAGAAAGAAAGAACCC 3529  
 QY 385 oGlnGlnGlyValTyrAsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSerGlnI 405  
 DB 3530 TCAGAAAGGCTCTTACATGAATGCAAGAAAGATGAATGGCGGAGCTTACAGTGAAT 3589  
 QY 405 eGlyThrLysGlyGluArgArgArgGlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSe 425  
 DB 3590 TGGATGAAAGGCGAGGCGCGGAGGGGCAAGAGGCGCAAGATGGCTTTTACAGGCTCTCAG 3649

Qy 425 rThralaThrlYsAspThrlYrAspAlaUhiSweGlnThrlLeuAlaProArg 443  
Db 3650 TACAGCCACCAAGAGACCTACGAGCCCTTCACTGACATGAGCCCTGCCCCCTCGC 3704

RESULT 3  
US-10-006-773-1.  
; Sequence 1, Application US/10006773  
; Publication No. US20020132983A1  
; GENERAL INFORMATION:  
; APPLICANT: Junghans, Richard P.  
; TITLE OF INVENTION: Antibodies as Chimeric Effector Cell Receptors Against Tumor Anti-  
; FILE REFERENCE: 003  
; CURRENT APPLICATION NUMBER: US/10/006,773  
; CURRENT FILING DATE: 2001-12-10  
; PRIOR APPLICATION NUMBER: 60/250,089  
; PRIOR FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 19  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 7654  
; TYPE: DNA  
; ORGANISM: Homo sapiens and Mus sp.  
; FEATURE:  
; NAME/KEY: CDS  
; LOCATION: (2428)..(3759)  
; OTHER INFORMATION: Chimeric IgTCR sequence contained in retroviral vector. Retrovi-  
; OTHER INFORMATION: al vector sequence (non-coding regions) are incidental to the inv-  
; OTHER INFORMATION: ention. The translated (coding region) is relevant to the invent-  
; OTHER INFORMATION: ion. (pertinent to Figure 3.)  
US-10-006-773-1

Alignment Scores:  
Pred. No.: 2.53e-95 Length: 7654  
Score: 1015.50 Matches: 225  
Percent Similarity: 57.74% Conservative: 25  
Best Local Similarity: 51.96% Mismatches: 58  
Query Match: 43.30% Indels: 125  
Gaps: 8

US-09-596-774-7 (1-443) x US-10-006-773-1 (1-7654)

Qy 131 G1yG1yG1ySerAspP1legInLeuThrG1nSerH1slySphLeuSerThrSerValG1y 150  
Db 2473 GGTGTTCACCTCCGACATCCAGCTGACCCGACCAAGCCAGCCAGCCGCGCGGTGGGT 2532

Qy 151 AspArgValSerIleThrCysAlaSerGlnAspValTyrAsnAlaValAlaTProTyr 170  
Db 2533 GACGAGTGAACCATCCTCTGTAAGCCAGTCAAGATGTGGTACTTCTGTAGCTTGGTAC 2592

Qy 171 GlnGlnLysProGlnGlnSerProLysLeuLeuIleTyrSerAlaSerSerArgTyrThr 190  
Db 2593 CAGCAGAAAGCCAGGTAAAGCTCCAAAGCTGCTGATCTAATGCAATCCACCCGCAACT 2652

Qy 191 G1yAlaProSerArgRhetHrGlnSerG1ySerG1yProAspPheThrPheThrIleSer 210  
Db 2653 GGTGTGCAAGCAATTCAGCGGTAGCGGTAGCGGTACCACTTCACCTTCCACCATAGC 2712

Qy 211 SerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGlnHis-----PheArgThr 228  
Db 2713 AGCTTCACGCAAGAGACATCGCACTACTACGCAAGATATAGCTCTATCGG--- 2769

Qy 229 ProPheThrPheG1ySerG1yThrlYsLeuGlnIleLys----- 241  
Db 2770 -----TCGTTCCGCAAGGAGCAAGGTGGAATCAAAAGAGGTGGCTCAGATCGGCT 2823

Qy 241 ----- 241  
Db 2824 GATCCGGCTCTGTGGCTCAGATCGAGCTCAACTGCTGTGAGACCGGTGAGGTGT 2883

Qy 242 -----AlaLeuGlnIleSerAsnSerVal----- 249  
Db 2884 GTGCAACTGGCCGGTCTCTGCGCTGTCTGCTCCGATCTGGCTTTCATTTCCACCACA 2943

Qy 249 ----- 249  
Db 2944 TATTGATGATGGTGGTGAAGAGCAGCAGCTGGAAAAGTCTTGATGTGATGGAAATT 3003

Qy 249 ----- 249  
Db 3004 CATCCAGATAGACATGATTAATCTATGCGCCGCTCTTAAGAGTATGATTTACATATCG 3063

Qy 249 ----- 249  
Db 3064 CGAGCAACGCCAAGAACACATGTTCTTCGAAATGACAGACCTTGAGACCCGAAACACC 3123

Qy 250 -----MetTyrPheSerSer 254  
Db 3124 GGGGTCTATTTTTGTGACACCTTTACTTGGCTTCCCTCGTTTGGTTATTTGGGCCAA 3183

Qy 255 ValValProValIleuGlnLysValAsnSerThrlYsProValIleuArgThrPro 274  
Db 3184 GGGACCCCGGTCAACCGTCCAGTGTGTAAGCCACACGACGACGCGCCGACCAACCA 3243

Qy 275 SerProValHisProThrG1yThrlYsSerGlnPro-----GlnArgProGlnAspCysArg 292  
Db 3244 ACAACCGCG---CCACCATCGCGTGCAGCCCTGTCTCTGCGCCACAGCGCGCTCGG 3300

Qy 293 ProArg-----GlySerValLysG1yThrlYsAspPheLeuGlnAspProLysIleu 310  
Db 3301 CCAACCGCGGGGGGCGAGTGCACAGAGGGGCTGACTTGCCTTGATCCCAACTC 3360

Qy 311 CysTyrIleuLeuAspG1yIleLeuPheIleTyrG1yValIleThrAlaLeuTyrIleu 330  
Db 3361 TGCTACCTGCTGATGAAATCCTTCTCATCTATGATGATCTACTGCTGCTTGTCCG 3420

Qy 331 ArgAlaIlePheSerArgSerAlaGlnThrAlaAlaAsnLeuGlnAspProAsnGlnLeu 350  
Db 3421 AGAGTGAAGTTCAGCAGAGCGCAGAGCCCGCCCTACAGAGCGCAGAACCGAGTTC 3480

Qy 351 TyrAsnGlnLeuAsnLeuGlnArgArgGlnGlnTyrTyrAspValIleuGlnLysIleArgAla 370  
Db 3481 TATTAAGAGCTCATCTTGAAGCAAGAGAGTACGATCTTTGGACAAAGACCTGGC 3540

Qy 371 ArgAspProGlnLeuG1yG1yLysGlnGlnArgArgAsnProGlnGlnIleValTyr 390  
Db 3541 CGGAGCCTGAGATGGGGGAGAAAG---CCGAGAAAGAGAAACCTTCAGAAAGGCTGTAC 3597

Qy 391 AsnAlaLeuGlnLysAspLysMetAlaGlnAlaTyrSerGlnIleG1yThrlYsG1yGln 410  
Db 3598 AATGAACCTGCAAGAAATAGATGGCGAGGCTTACAGTGAATGGATGGAGAAAGCGAG 3657

Qy 411 ArgArgArgG1yLysG1yHisAspG1yLeuTyrGlnG1yLeuSerThrlYsAsp 430  
Db 3658 CGCCGAGAGGGGCAAGGGGACCATGGCTTTACAGAGGTCTCAGTACAGCCACCAAGAGC 3717

Qy 431 ThrTyrAspAlaLeuHisMetGlnThrlYsAlaProArg 443  
Db 3718 ACCTAGCAGCGCTTTCACATGAGCGCCCTCCCTCGC 3756

RESULT 4  
US-10-006-771A-1  
; Sequence 1, Application US/10006771A  
; Publication No. US20020165360A1  
; GENERAL INFORMATION:  
; APPLICANT: Junghans, Richard P.  
; TITLE OF INVENTION: Chimeric Effector Cell Receptors Against Carcinoembryonic Antigen  
; FILE REFERENCE: 002  
; CURRENT APPLICATION NUMBER: US/10/006,771A  
; CURRENT FILING DATE: 2002-06-04  
; PRIOR APPLICATION NUMBER: 60/250,090  
; PRIOR FILING DATE: 2000-11-30  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 1  
; LENGTH: 7654  
; TYPE: DNA



```

Db      1 CAGGTGAAGCTGCAGACAGTCCGAGTTGAAGAACCGGGGTGAGACCTCAAGATC 60
Qy      21 SerCyslysaIaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValysGlnaIa 40
Db      61 AGCTCAAGAGGCTTCGCTTATACCTTCAACCGGTGGTATGAACCTGGGTGAAGCAGGCT 120
Qy      41 ProGlyGlnGlyLeuLeuTyrPheMetGlyTyrPheLeuAsnThrSerThrGlyGlnSerThrPhe 60
Db      121 CCGGGCAAGGCTTTAAAGTGAAGTGGCTGGATTAAACCAAACTGGTGAAGCAACCTAT 180
Qy      61 AlaSerPhePheIysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db      181 GTTGAAGAGTTTAAAGGTGCTTCCCTTCTTTGGAGACCTCTGCCACACACCTCCAT 240
Qy      81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGln 100
Db      241 TTGCAGATCAACACCTCAAAATGAGGACACGGCTAAATATTTCTGTGACAGTTGGGAC 300
Qy      101 ValTyrHisGlyTyrVal-----ProTyrTyrGlyGlnGlyThrValThrVal 117
Db      301 TTCTAT---GATTACGTGAAGCTATGGAATTACTGGGGCCAAAGGACACGCTACCGCTC 357
Qy      118 Ser-----SerGlyGlyGlySerGly 125
Db      358 TCCAAGATCTCGTGGCGGCTCGGCTCGGGGTGGTGGGTGGGTGGGTGGGTGGGTGGGT 417
Qy      126 GlyGlyGlySerGlyGlyGlyGly---SerAspIleGlnLeuThrGlnSerHisLysPhe 144
Db      418 GGTGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 477
Qy      145 LeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSerGlnAspValTyr 164
Db      478 ATGTCCACTTCAGTAGTGAATCGTGTGACGCTCAGCTGCAAGCAGACGATGTGGCT 537
Qy      165 AsnAlaValAlaTyrGlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSer 184
Db      538 ACGAATGTTGCTGGTATCAACAGAAACCGGGTCAATCCCGAAAGCACTGATTACTCG 597
Qy      185 AlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGlyProAsp 204
Db      598 GCATCTACCTGATCAGTGAAGTGGTCCCGATGCTTACCCGCAAGTGGTTCGGGACCGAT 657
Qy      205 PheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGln 224
Db      658 TTCAAGCTCAACCATCAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 717
Qy      225 HisPheArgThrPro---PheThrPheGlySerGlyThrLysLeuGlnIleLys 241
Db      718 TATTACACCTATCCGTTATTCACTGCGGCTCGGGGACCAAGTTGGAATGAAG 771

```

## RESULT 6

```

US-10-150-762-38
; Sequence 38, Application US/10150762
; Publication No. US20030103948A1
; GENERAL INFORMATION:
; APPLICANT: Goshorn, Stephen C.
; APPLICANT: Graves, Scott S.
; APPLICANT: Schultz, Joanne E.
; APPLICANT: Lin, Yukang
; APPLICANT: Sanderson, James A.
; APPLICANT: Reno, John M.
; APPLICANT: Dearstyne, Erica A.
; TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND
; FILE REFERENCE: 690022.547C2
; CURRENT APPLICATION NUMBER: US/10/150.762
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 38
; LENGTH: 771
; TYPE: DNA

```

```

; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: PRIA3 single chain antibody-genomic Streptavidin
; US-10-150-762-38

```

## Alignment Scores:

```

Pred. No.: 2.04e-94 Length: 771
Score: 994.00 Matches: 193
Percent Similarity: 82.95% Conservative: 21
Best Local Similarity: 74.81% Mismatches: 26
Query Match: 42.39% Indels: 18
DB: 15 Gaps: 5

```

US-09-596-774-7 (1-443) x US-10-150-762-38 (1-771)

```

Qy      1 GlnIleGlnLeuValGlnSerGlyProGlnLeuLysProGlyGlnThrValIle 20
Db      1 CAGGTGAAGCTGCAGACAGTCCGAGTTGAAGAACCGGGGTGAGACCTCAAGATC 60
Qy      21 SerCyslysaIaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValysGlnaIa 40
Db      61 AGCTCAAGAGGCTTCGCTTATACCTTCAACCGGTGGTATGAACCTGGGTGAAGCAGGCT 120
Qy      41 ProGlyGlnGlyLeuLeuTyrPheMetGlyTyrPheLeuAsnThrSerThrGlyGlnSerThrPhe 60
Db      121 CCGGGCAAGGCTTTAAAGTGAAGTGGCTGGATTAAACCAAACTGGTGAAGCAACCTAT 180
Qy      61 AlaSerPhePheIysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db      181 GTTGAAGAGTTTAAAGGTGCTTCCCTTCTTTGGAGACCTCTGCCACACACCTCCAT 240
Qy      81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGln 100
Db      241 TTGCAGATCAACACCTCAAAATGAGGACACGGCTAAATATTTCTGTGACAGTTGGGAC 300
Qy      101 ValTyrHisGlyTyrVal-----ProTyrTyrGlyGlnGlyThrValThrVal 117
Db      301 TTCTAT---GATTACGTGAAGCTATGGAATTACTGGGGCCAAAGGACACGCTACCGCTC 357
Qy      118 Ser-----SerGlyGlyGlySerGly 125
Db      358 TCCAAGATCTCGTGGCGGCTCGGCTCGGGGTGGTGGGTGGGTGGGTGGGTGGGTGGGT 417
Qy      126 GlyGlyGlySerGlyGlyGlyGly---SerAspIleGlnLeuThrGlnSerHisLysPhe 144
Db      418 GGTGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGT 477
Qy      145 LeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSerGlnAspValTyr 164
Db      478 ATGTCCACTTCAGTAGTGAATCGTGTGACGCTCAGCTGCAAGCAGACGATGTGGCT 537
Qy      165 AsnAlaValAlaTyrGlnGlnLysProGlyGlnSerProLysLeuLeuIleTyrSer 184
Db      538 ACGAATGTTGCTGGTATCAACAGAAACCGGGTCAATCCCGAAAGCACTGATTACTCG 597
Qy      185 AlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlySerGlyProAsp 204
Db      598 GCATCTACCTGATCAGTGAAGTGGTCCCGATGCTTACCCGCAAGTGGTTCGGGACCGAT 657
Qy      205 PheThrPheThrIleSerSerValGlnAlaGlnAspLeuAlaValTyrPheCysGlnGln 224
Db      658 TTCAAGCTCAACCATCAGCAATGATGATGATGATGATGATGATGATGATGATGATGAT 717
Qy      225 HisPheArgThrPro---PheThrPheGlySerGlyThrLysLeuGlnIleLys 241
Db      718 TATTACACCTATCCGTTATTCACTGCGGCTCGGGGACCAAGTTGGAATGAAG 771

```

## RESULT 7

```

US-10-244-821-38
; Sequence 38, Application US/10244821
; Publication No. US2003014323A1
; GENERAL INFORMATION:

```

APPLICANT: Goshorn, Stephen Charles  
 APPLICANT: Graves, Scott Scoll  
 APPLICANT: Schultz, Joanne Elaine  
 APPLICANT: Lin, Yukang  
 APPLICANT: Sanderson, James Allen  
 APPLICANT: Reno, John M.  
 APPLICANT: Dearstyne, Erica A.  
 TITLE OF INVENTION: STREPTAVIDIN EXPRESSED GENE FUSIONS AND  
 TITLE OF INVENTION: METHODS OF USE THEREOF  
 FILE REFERENCE: 690022.547C3  
 CURRENT APPLICATION NUMBER: US/10/244,821  
 CURRENT FILING DATE: 2002-09-16  
 NUMBER OF SEQ ID NOS: 92  
 SOFTWARE: FASTSEQ for Windows Version 4.0  
 SEQ ID NO 38  
 LENGTH: 771  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURES:  
 OTHER INFORMATION: PRLA3 single chain antibody-genomic streptavidin  
 OTHER INFORMATION: fusion construct  
 US-10-244-821-38

Alignment Scores:  
 Pred. No.: 2,04e-94 Length: 771  
 Score: 994.00 Matches: 193  
 Percent Similarity: 82.95% Conservative: 21  
 Best Local Similarity: 74.81% Mismatches: 26  
 Query Match: 42.39% Indels: 18  
 DB: Gaps: 5

US-09-596-774-7 (1-443) x US-10-244-821-38 (1-771)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLeuProGlyGlnThrValLysIle 20  
 Db 1 CAGGTGAAGCTGCGAGCTGCTCCGAGTTGAAGACCGGGTGAACCGTCAAGATC 60  
 QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40  
 Db 61 AGCTGCAAGGCTTCTGTTATACCTTCACCGTGTGATGATGAGTGAACGAGCT 120  
 QY 41 ProGlyGlnGlyLeuLysTyrPheMetGlyTyrIleAsnThrSerThrGlyLysSerThrPhe 60  
 Db 121 CCGGGCAAGGCTTTAACTGATGGATGGATTAACCAACAACTGGTGAAGCAACTAT 180  
 QY 61 AlaPhePhePheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80  
 Db 181 GTTGAAGAGTTTAAAGGCTGCTTTCCTTCTTGGAGACCTGACACCACTGCTAT 240  
 QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100  
 Db 241 TTGCAAGATCAACCACTCAAAATGAGGACAGGCTTAATATTTCTGACAGTTGGAC 300  
 QY 101 ValTyrHisGlyTyrVal-----ProTyrTrpGlyGlnGlyThrThrValThrVal 117  
 Db 301 TTCTAT---GATTACGTGAAGCTATGATTACTGGGGCCAAAGGACCAAGCTACCGTC 357  
 QY 118 Ser-----SerGlyGlyGlySerGly 125  
 Db 358 TCCAAAGATCTGTGGGCGGTGGCTCGGGCGGTGGTGGTGGGCGGCGCTCGGCT 417  
 QY 126 GlyGlyGlySerGlyGlyGlyGly-----SerAspIleGlnLeuThrGlnSerHisLysPhe 144  
 Db 418 GGTGTGTGGTGGGCGGCGGCGGCTCGAGCATATTTGAATGACCACTTCACAGTTTC 477  
 QY 145 LeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSerGlnAspValTyr 164  
 Db 478 ATGTCACATTCAGTAGTGATGATGCTGACGCTCACTCAACAGCAAGTATGGGT 537  
 QY 165 AsnAlaValAlaTyrTyrGlnGlnLysPheProGlyGlnSerProLysLeuLeuIleTyrSer 184  
 Db 538 ACGAATGTTGCTGGTATCAACAGAAACCGGGTCAATCCCGAAAGCACTGATTACTCG 597

QY 185 AlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGlyProAsp 204  
 Db 598 GCATCTCAACCTTACAGTGGTGTCCGGATGCTTACACCGGACAGTGTTCGGACCAT 657  
 QY 205 PheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyrPheCysGlnGln 224  
 Db 658 TTCAGGCTCAACCAACAGCAAGTACAGTCTGACAGTCTGGGAGTATTTCTGCATCAA 717  
 QY 225 HisPheArgThrPro---PheThrPheGlySerGlyTyrHisLeuGluIleLys 241  
 Db 718 TATTACACTTACCGTTATTCAGTCTGGGAGCAACAGTTGGAATGAAAG 771

RESULT 8

US-09-887-853-1  
 Sequence 1, Application US/09887853  
 Patent No. US20020168375A1  
 GENERAL INFORMATION:  
 APPLICANT: Huston, James S.  
 Oppermann, Hermann  
 Houston, L. L.  
 Ring, David B.  
 TITLE OF INVENTION: Biosynthetic Binding Proteins For  
 Imaging

NUMBER OF SEQUENCES: 11  
 CORRESPONDENCE ADDRESS:  
 ADDRESSER: Teasta, Hurwitz & Thibault/Parent Department  
 STREET: Exchange Place, 53 State Street  
 CITY: Boston  
 STATE: Massachusetts  
 COUNTRY: USA  
 ZIP: 02109

COMPUTER READABLE FORM:  
 MEDIUM TYPE: floppy disk  
 COMPUTER: IBM PC compatible  
 OPERATING SYSTEM: PC-DOS/MS-DOS  
 SOFTWARE: Patent Release #1.0, Version #1.25  
 CURRENT APPLICATION DATA:  
 APPLICATION NUMBER: US/09/887,853  
 FILING DATE: 21-Jun-2001  
 CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:  
 APPLICATION NUMBER: US/08/133,804  
 FILING DATE: <Unknown>  
 ATTORNEY/AGENT INFORMATION:  
 NAME: Kelley, Robin D.  
 REGISTRATION NUMBER: 34,637  
 REFERENCE/DOCKET NUMBER: 2054/22  
 TELECOMMUNICATION INFORMATION:  
 TELEPHONE: 617-248-7477  
 TELEFAX: 617-248-7100

INFORMATION FOR SEQ ID NO: 1:  
 SEQUENCE CHARACTERISTICS:  
 LENGTH: 909 base pairs  
 TYPE: nucleic acid  
 STRANDEDNESS: single  
 TOPOLOGY: linear  
 MOLECULE TYPE: cDNA

FEATURE:  
 NAME/KEY: CDS  
 LOCATION: 3..752  
 OTHER INFORMATION: /product= "741F8 sfv" C-terminal  
 Gly4-Cys"  
 SEQUENCE DESCRIPTION: SEQ ID NO: 1:  
 US-09-887-853-1

Alignment Scores:  
 Pred. No.: 5.23e-93 Length: 909  
 Score: 981.50 Matches: 185  
 Percent Similarity: 85.89% Conservative: 22  
 Best Local Similarity: 76.76% Mismatches: 33  
 Query Match: 41.86% Indels: 1  
 DB: Gaps: 1

US-09-596-774-7 (1-443) x US-09-887-853-1 (1-909)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGluGluThrValLysIle 20
DB 9 GAGATCCAAATGGTGCAGTCTGACCTGAGCTGAAGAACCTCGAGAGACGTCAGATC 68
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
DB 69 TCCTGCAAGGCTTCTGGGTATACCTTCAACAACTATGAGATGAATCGGGTGAAGCAGCT 128
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrTrpIleAsnThrSerThrGlyLysSerThrPhe 60
DB 129 CAGGAAGAGGTTTAAAGTGAGTGAGGCTGAGTAAACACCACTGAGAGACCAACATAT 188
QY 61 AlaAspAspPheLysGlyTyrPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 189 GCTGAAGAGTTCAACGAGCAGGTTTCCCTCTTCTTGGAACTCTGCCAGACATGCCAT 248
QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
DB 249 TTGCAGATCAACAACCTCAAAAATGAGGACACGCGCTACATATTCTGTGAAGCAATTT 308
QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
DB 309 ATTACCTAAGCGGGGTTCTTCTAAGCTGAGGCTGAGCTGAGCTGAGCTGAGCTGAG 365
QY 121 GlyGlyLysSerGlyGlyGlySerGlySerGlySerAspIleGlnLeuThrGln 140
DB 366 TCGAGCTCCTCGGATCTTCTATAGCGGTTCCAGCTGAGCTGAGCTGAGCTGAGCTGAG 425
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
DB 426 TCTCTTAATTCATGTCACGCTGAGTGAGGACAGGCTCAGCATCTCTCGAAGGCGAGT 485
QY 161 GlnAspValTyrAsnAlaValAlaIleTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB 486 CAGAGTGTGAGTACTGCTGTCGCTGATGATCAACAAACAGGACATCTCTTAACTA 545
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 546 CTGATTTACTGACATCCACCGGACACCTGAGTCCCTGATGCTTCAACGAGGAGGA 605
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 606 TCTGGGACAGATTATCTTCAACCATCAGCAGTGTGAGGCTGAAGACCTGGCACTTCAT 665
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrHisLeuGluIle 240
DB 666 TACTGTACGACAACTTATAGAGTCCCTGACGTTCCGAGGGGAGCAACGCTGGAGATA 725
QY 241 Lys 241
DB 726 AAA 728

```

# RESULT 9

```

US-10-683-547-1
; Sequence 1, Application US/10683547
; Publication No. US20050058638A1
; GENERAL INFORMATION:
; APPLICANT: Huseon, J.
; APPLICANT: Houston, L.L.
; APPLICANT: Ring, D.
; APPLICANT: Oppermann, H.
; TITLE OF INVENTION: BIOSYNTHETIC BINDING PROTEINS FOR IMMUNO-TARGETING
; FILE REFERENCE: CIBT-P01-130
; CURRENT APPLICATION NUMBER: US/10/683,547
; PRIOR FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: US/09/558,741
; PRIOR FILING DATE: 2000-04-26
; PRIOR APPLICATION NUMBER: 07/831,967
; PRIOR FILING DATE: 1992-02-06
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn version 3.1

```

```

; SEQ ID NO 1
; LENGTH: 909
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: 741P8 sFv'
; NAME/KEY: CDS
; LOCATION: (3) -(752)
; OTHER INFORMATION:
US-10-683-547-1

```

## Alignment Scores:

Pred. No.:	Score:	Length:	Matches:	Conservative:	Mismatches:	Indels:	Gaps:
9.52e-92	969.50	183	22	35	1		
Percent Similarity:	85.06%						
Best Local Similarity:	75.93%						
Query Match:	41.34%						

US-09-596-774-7 (1-443) x US-10-683-547-1 (1-909)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysLysProGluGluThrValLysIle 20
DB 9 GAGATCCAAATGGTGCAGTCTGACCTGAGCTGAAGAACCTCGAGAGACGTCAGATC 68
QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValLysGlnAla 40
DB 69 TCCTGCAAGGCTTCTGGGTATACCTTCAACAACTATGAGATGAATCGGGTGAAGCAGCT 128
QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrTrpIleAsnThrSerThrGlyLysSerThrPhe 60
DB 129 CAGGAAGAGGTTTAAAGTGAGTGAGGCTGAGTAAACACCACTGAGAGACCAACATAT 188
QY 61 AlaAspAspPheLysGlyTyrPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
DB 189 GCTGAAGAGTTCAACGAGCAGGTTTCCCTCTTCTTGGAACTCTGCCAGACCTGCTAT 248
QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
DB 426 TCTCTTAATTCATGTCACGCTGAGTGAGGACAGGCTCAGCATCTCTCGAAGGCGAGT 485
QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
DB 309 ATTACCTAAGCGGGGTTCTTCTAAGCTGAGGCTGAGCTGAGCTGAGCTGAGCTGAG 365
QY 121 GlyGlyLysSerGlyGlyGlySerGlySerGlySerAspIleGlnLeuThrGln 140
DB 366 TCGAGCTCCTCGGATCTTCTATAGCGGTTCCAGCTGAGCTGAGCTGAGCTGAGCTGAG 425
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
DB 426 TCTCTTAATTCATGTCACGCTGAGTGAGGACAGGCTGAGCTGAGCTGAGCTGAGCTGAG 485
QY 161 GlnAspValTyrAsnAlaValAlaIleTyrGlnGlnLysProGlyGlnSerProLysLeu 180
DB 486 CAGAGTGTGAGTACTGCTGTCGCTGATGATCAACAAACAGGACATCTCTTAACTA 545
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
DB 546 CTGATTTACTGACATCCACCGGACACCTGAGTCCCTGATGCTTCAACGAGGAGGA 605
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
DB 606 TCTGGGACAGATTATCTTCAACCATCAGCAGTGTGAGGCTGAAGACCTGGCACTTCAT 665
QY 221 PheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyTyrHisLeuGluIle 240
DB 666 TACTGTACGACAACTTATAGAGTCCCTGACGTTCCGAGGGGAGCAACGCTGGAGATA 725
QY 241 Lys 241
DB 726 AAA 728

```

## RESULT 10

US-08-812-393A-1

Sequence 1, Application US/08812393A

Publication No. US2001007152A1

GENERAL INFORMATION:

APPLICANT: SHERMAN, Linda A.

APPLICANT: LUSTGARTEN, Joseph

TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING

TITLE OF INVENTION: T CELL RECEPTORS SPECIFIC FOR HUMAN HLA-RESTRICTED TUMOR

TITLE OF INVENTION: ANTIGENS

NUMBER OF SEQUENCES: 64

CORRESPONDENCE ADDRESSES:

ADDRESS: MORRISON &amp; FOERSTER

STREET: 2000 Pennsylvania Avenue, NW, suite 5500

CITY: Washington

STATE: DC

COUNTRY: USA

ZIP: 20006-1888

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/812.393A

FILING DATE: 05-MAR-1997

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Murashige, Kate H

REGISTRATION NUMBER: 29,959

REFERENCE/DOCKET NUMBER: 31333-20001.00

TELECOMMUNICATION INFORMATION:

TELEPHONE: 202-887-1500

TELEFAX: 202-822-0168

TELEX:

INFORMATION FOR SEQ ID NO: 1:

SEQUENCE CHARACTERISTICS:

LENGTH: 1350 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: linear

FEATURE:

NAME/KEY: Coding Sequence

LOCATION: 1...1332

OTHER INFORMATION:

US-08-812-393A-1

US-09-596-774-7 (1-443) x US-08-812-393A-1 (1-1350)

Alignment Scores:

Pred. No.: 2,28e-87

Score: 930.00

Percent Similarity: 59.33%

Best Local Similarity: 48.22%

Query Match: 39.66%

DB: 8

Gaps: 14

Length: 1350

Matches: 217

Conservative: 50

Mismatches: 145

Indels: 38

Gaps: 14

US-09-596-774-7 (1-443) x US-08-812-393A-1 (1-1350)

QY 3 GlnLeuValGlnSerGlyProGlu-----LeuIysLysProGlyGluThrValLysLeu 20

DB 70 CAGCAAGGAGCAGAGAGTCCCGCATCTTGCTTGCAGAGGAGAGACGACAGCTC 129

QY 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40

DB 130 CAGGTGAGCTTTTC-----ATCTTACAAACAG---GTGACGTGTTTACCAACT 180

QY 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrPheAsnThrSerThrGlyGluSerThrPhe 60

DB 181 CCGGGGAGAGATCGTCAGCCTGTGTGACATCTTCTGGAGACAAAGCAGAGT----- 234

QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80

DB 235 -----GGGAGAGCTGACATCCACAGACATTAAGAAAGTCGAGCTCT 279

QY 81 LeuGlnIleAsnAsnLeuLysSerGluAspMetAlaThrTyrPheCysAlaArgTyrGlu 100

DB 280 TTGCACATTTCTCTCTCCACATCATCAGACTCAGACCTATTCCTGTGCTCAATTC 339

QY 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerSerGly 120

DB 340 GGAGAAAGCAATGCAAGCTTACCTTGGGAAAGCACTAAACTTCGTGTAATTCAGT 399

QY 121 GlyGlyLysSerGlyGlyGlyGlySerGlyGlyLysSerAspIleGlnLeuThrGln 140

DB 400 GCGGAGGCTGTGCGGGGCGTGGATCCGGGGGTGAGGCTCAGAGGCTCAGACCCAA 459

QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160

DB 460 AGCCCAAGAAACAGGTGCGAGTAACAGAGAAAGGTGACATTGAGCTGTATCAGACT 519

QY 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180

DB 520 AATAAC---CACAACAATGATCTGATCGGAGACAGACGGGCGATGGCTGAGCTG 576

QY 181 LeuIleTyrSer-----AlaSerSerArgTyrThrGlyValProSerArgPheThr 197

DB 577 ATCCATTATTCATATGCTGCTGCGAGCATGGAAGAGATATCCCTGATGATGATCAAG 636

QY 198 GlySer---GlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 216

DB 637 GCCTCCAGACCAACCAAGAGAACTCTCCCTCATCTGAGATTGAGTACCCCTCAG 696

QY 217 LeuAlaValTyrPheCys-----GlnGlnHisPheArgThrProPheThrPheGly 233

DB 697 ACATCAGGTACTTCTGTGCGAGCGGTGAGACAGAGACCAAGAAATTAATTTTCGCT 756

QY 234 SerGlyThrLysLysGluIleLysAlaLeuGluIleSerAsnSerValMetTyrPheSer 253

DB 757 CATTGAAACCAAGCTGTCTGCTGCTGACT-----AGTAATCTCATCATGATTTACG 807

QY 254 SerValValProValLeuGlnLysValAsnSerThrThrLysProValLeuArgThr 273

DB 808 CACTTCGTGCGGCTTCTGCTGCGAGCCAGAGCCCAAGAGC---CCAGCGCCCGCA 864

QY 274 ProSerProValHisProThrGlyThrSerGlnProGlnArgProGluAspCysArgPro 293

DB 865 CCAACACCGCGG---CCACCATGCGGTGCGAGCCCTG-----TCCCTGCGCCA 912

QY 294 ArgGlySerValLysGlyThrGlyLeuAspPheLeuGluAspProLysLeuCysTyrLeu 313

DB 913 TCTAGTCT-----AGAGATCCCAACTGCTTACTCT 945

QY 314 LeuAspGlyIleLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLys 333

DB 946 CTGATGGAATCTCTTCACTATGATGATGATCTTCACTGCTGCTGCTGATGAGTGAAG 1005

QY 334 PheSerArgSerAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnIleTyrAsnGlu 353

DB 1006 TTCAGCAGAGCGAGAGCGCCCGCGTACAGAGGCGCAAGACAGCTTAAACAGAG 1065

QY 354 LeuAsnLeuGlyArgGluGluTyrAspValLeuGlnLysLysValAspAlaArgAspPro 373

DB 1066 CTCAATCTTAGAGCAAGAGAGAGATGATGATTTTGAACAAGAGACGTGGCGGAGCCT 1125

QY 374 GluMetGlyGlyLysGlnGlnArgArgAsnProGlnGlnGlyValTyrAsnAlaLeu 393

DB 1126 GAGATGGGGGAAAG---CCGAGAGAGAGAACCTCAGAGAGCGCTGTAACAAAGACTG 1182

QY 394 GlnLysAspLysMetAlaGluAlaTyrSerGluIleIleTyrLysGlyGluArgArgArg 413

DB 1183 CAGAAATATAGATGGGAGGCGCTTACAGTGAATTTGGATGAAGAGCGCGGAGAG 1242

QY 414 GlyLysGlyHisAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAsp 433

Db 1243 GGCAAGGGGACGAGTGGCTTTACCAAGGCTCTCAATGACGACCAAGAGCACTTACAC 1302  
Qy 434 AAlaLeuHleuMeGlnThrLeuAlaProArg 443  
Db 1303 GCCCTTCAATGACAGGCGCTGCCCCCTCGC 1332

RESULT 11  
US-09-774-681-1  
; Sequence 1, Application US/09774681  
; Publication No. US20030208780A1  
; GENERAL INFORMATION:  
; APPLICANT: Sunol Molecular Corporation  
; APPLICANT: Sherman, Linda  
; APPLICANT: Lustgarten, Joseph  
; TITLE OF INVENTION: RECOMBINANT CONSTRUCTS ENCODING T CELL  
; FILE REFERENCE: 31333-20001.01  
; CURRENT APPLICATION NUMBER: US/09/774/681  
; PRIOR FILING DATE: 2001-02-01  
; PRIOR APPLICATION NUMBER: US 08/812,393  
; PRIOR FILING DATE: 1997-03-05  
; PRIOR APPLICATION NUMBER: US 60/012,845  
; NUMBER OF SEQ ID NOS: 65  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 1  
; LENGTH: 1350  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: Polynucleotide derivative of effective T cell  
; NAME/KEY: CDS  
; LOCATION: (1)...(1350)  
US-09-774-681-1

Alignment Scores:  
Pred. No.: 3.7e-87 Length: 1350  
Score: 928.00 Matches: 217  
Percent Similarity: 59.33% Conservative: 50  
Best Local Similarity: 48.22% Mismatches: 145  
Query Match: 39.57% Indels: 38  
Gaps: 14

US-09-596-774-7 (1-443) x US-09-774-681-1 (1-1350)

Qy 3 GlnLeuValGlnSerGlyProGlu-----LeuLysProGlyGluThrValLysIle 20  
Db 70 CAGCAAGTGCAGCAGAGTCCCGCATCTCTGTCGACGAGGCGGAGAAAGCAGAGCTC 129

Qy 21 SerCysLysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTyrValLysGlnAla 40  
Db 130 CAGGTGAGCTTTCC-----ATCTTTACAAACAG--GTGAGGTGTTTACCAAGT 180

Qy 41 ProGlyGlnGlyLeuLysTyrMetGlyTyrPheAsnThrSerThrGlyLysThrPhe 60  
Db 181 CCTGGGGAGAGACTCGTCAGCCTGTGTGACAAATCCTTGGGAGAAAGCAGAGT----- 234

Qy 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80  
Db 235 -----GGGAGAGCTGACATCCACCAACAGTCAATTAAGAAAGCTCGAGCTCT 279

Qy 81 LeuGlnIleAsnLeuLysSerGlyAspMetAlaThrTyrPheCysAlaArgTyrGlu 100  
Db 280 TTGACACATTTCTCTCTCCGATCAGACTCAGGCACTTAATCTCTGTGCTCAATTTCT 339

Qy 101 ValTyrHisGlyTyrValProTyrTyrGlyGlnGlyThrThrValThrValSerGly 120  
Db 340 GGAAGAGCAATGCAAGAGTAACTTGGGAAAGGCACTAACTCTCTGTTAAATACAGT 399

Qy 121 GlyGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 140  
Db 121 GlyGlyGlySerGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGlyGly 140

Db 400 GCGGAGGCTGTGGCGGGGGGATCCGAGGAGTGAAGCTCAAGAGCTGACATACCCA 459  
Qy 141 SerHisPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160  
Db 460 AGCCCAAGAAACAAGGTGACGATACAGAGAGAAAGGTGACATTAAGCTTAATCAGACT 519

Qy 161 GlnAspValTyrAsnAlaValAlaTyrTyrGlnGlnLysProGlyGlnSerProLysLeu 180  
Db 520 AATAAC--CAACAACAATGATCTGTATTCGGGAGACAGGCGGATGGGCTGAGGCTG 576

Qy 181 LeuIleTyrSer-----AlaSerSerArgTyrThrGlyValProSerArgPheThr 197  
Db 577 ATCATTAATTCATATGTGTGTGGAGAGACTGAGAAAGAAATATTCCTGATGATACAG 636

Qy 198 GlySer--GlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAsp 216  
Db 637 GCCCTCAGACCAAGCCAAAGAAAGAACTCTCCCTCATTTGTGAGTGGTATCCCTCTCAG 696

Qy 217 LeuAlaValTyrPheCys-----GlnGlnHisPheArgThrProPheThrPheGly 233  
Db 697 ACATCACTGATCTCTGTGTGCGCAGCGGTGAGACAGGACCAAGAAAGATTATTTTCGT 756

Qy 234 SerGlyThrLysLeuGlnIleLysAlaLeuGlnIleSerAsnSerValMetTyrPheSer 253  
Db 757 CATGGAACCAAGCTGTCTGTCTGACT-----AGTAATCTCATGATGATCTTCAGC 807

Qy 254 SerValValProValLeuGlnLysValAsnSerThrThrLysProValLeuArgThr 273  
Db 808 CACTTCGTGCGGTCTCTCTGCGCAGGAAAGCCACACACGACG--CCAGGCGCCGACACA 864

Qy 274 ProSerProValHisProThrGlyThrSerGlnProGlnArgProGluAspCysArgPro 293  
Db 865 CCAACACCGCG--CCACCATCGCGTGCACCGCTCG-----TCCCTGCGCCCA 912

Qy 294 ArgGlySerValLysGlyThrGlyLysLeuAspPheLeuGluAspProLysLeuGlyLeu 313  
Db 913 TCTAGTCT-----AGAGATCCCAACTGCTGACTCG 945

Qy 314 LeuAspGlyLysLeuPheIleTyrGlyValIleIleThrAlaLeuTyrLeuArgAlaLys 333  
Db 946 CTGATGAGATCTCTTCATCTATGATGTCATTCATTCATTCATTCATTCATTCATTCAT 1005

Qy 334 PheSerArgSerAlaGluThrAlaAlaAsnLeuGluAspProGlnLysThrGln 353  
Db 1006 TTCAGCAGAGCGCAGACCGCCCGCTACACAGAGGCGCAGACCGCTCTTAAACGAG 1065

Qy 354 LeuAsnLeuGlyArgArgGluGluTyrAspValLeuGlnLysLysArgAlaArgAspPro 373  
Db 1066 CTCATCTCTAGACGAAGAGAGAGTACGATGTTTGAACAAGAGTGGCGGAGCCCT 1125

Qy 374 GluMetGlyLysGlnGlnArgArgAspProGlnGlyValTyrAsnAlaLeu 393  
Db 1126 GAGATGGGGGAGAAAG--CCGAGAGAGAAAGAAAGCCCTCAGAGAGGCTGTACATGAATG 1182

Qy 394 GlnLysAspLysMetAlaGluAlaTyrSerGluIleGlyThrLysGlyGluLysArgArg 413  
Db 1183 CAGAAAGATTAAGTGGCGGAGGCTTACAGTGAATTTGGATGAAGAGCAGCGCGGAGG 1242

Qy 414 GlyLysGlyHisAspGlyLeuTyrGlnGlyLysSerThrAlaThrLysAspThrTyrAsp 433  
Db 1243 GGCAGAGGGGACAGATGGCTTTACAGAGGATCTCAGTACAGCAGCAGCAAGGACCTACGAGC 1302

Qy 434 AlaLeuHisMetGlnThrLeuAlaProArg 443  
Db 1303 GCCCTTCAATGACAGGCGCTGCCCCCTCGC 1332

RESULT 12  
US-10-239-656-78  
; Sequence 78, Application US/10239656  
; Publication No. US20040038339A1  
; GENERAL INFORMATION:  
; APPLICANT: KUFER, PETER  
; APPLICANT: RIETHMULLER, GERT



APPLICANT: LUTTERBUSE, RALF  
 APPLICANT: BORSCHERT, KATRIN  
 APPLICANT: KISCHEL, ROMAN  
 APPLICANT: MAYER, MONIKA  
 APPLICANT: HORNEISTER, ROBERT  
 TITLE OF INVENTION: MULTIFUNCTIONAL POLYPEPTIDES COMPRISING A BINDING SITE  
 TITLE OF INVENTION: TO AN EPITOPE OF THE NKGD RECEPTOR COMPLEX  
 FILE REFERENCE: 029976/0106  
 CURRENT APPLICATION NUMBER: US/10/239,656  
 PRIOR APPLICATION NUMBER: PCT/EP01/03414  
 PRIOR FILING DATE: 2001-03-26  
 PRIOR APPLICATION NUMBER: EP 00106467.4  
 PRIOR FILING DATE: 2000-03-24  
 NUMBER OF SEQ ID NOS: 92  
 SOFTWARE: PatentIn Ver. 2.1  
 SEQ ID NO: 78  
 LENGTH: 1515  
 TYPE: DNA  
 ORGANISM: Artificial Sequence  
 FEATURE:  
 OTHER INFORMATION: Description of Artificial Sequence: Synthetic 3810xps-  
 OTHER INFORMATION: 23 bispecific single chain fv  
 US-10-239-656-78

Alignment Scores:

Pred. No.: 5 476-86 Length: 1515  
 Score: 917.50 Matches: 185  
 Percent Similarity: 57.38% Conservative: 25  
 Best Local Similarity: 50.55% Mismatches: 31  
 Query Match: 39.13% Indels: 125  
 DB: 17 Gaps: 4

US-09-596-774-7 (1-443) x US-10-239-656-78 (1-1515)

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuLysProGluGluThrValLysIle 20  
 DB 382 CAGGTGCAACTGACGACGACGAGCGCTGAGTGAAGAAAGCCGAGAGACAGTCAAGATC 441  
 QY 21 SerCysblyValAspSerGlyProPheThrAspGlyMetAspTrpValLysGlnIle 40  
 DB 442 TCCTGCAAGGCTTCGGGTATACCTTCACAACTGTGAATGAACCTGGGTGAAGCAGGCT 501  
 QY 41 ProGluGlnGlyLeuValTrpMetGlyTrpIleAspThrSerThrGlyGluSerThrPhe 60  
 DB 502 CCAAGAAAGGCTTCAAGTGGAGGCTGATAAACCTACACTGAGAGCCAAACATATC 561  
 QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAspThrAlaTrp 80  
 DB 562 GGTATGACTTCAGAGGAGCAGGCTTCTCTTGGAAACCTGCGCAGCAGCTGCTAT 621  
 QY 81 LeuGlnIleAspAspLeuLysSerGluAspMetAlaThrTrpPheCysAlaArg----- 98  
 DB 622 TTGCAGATCAACAACCTCAAAAATGAGACACGCTACATATTTCTGTCAAGATTCAAC 681  
 QY 98 ----- 98  
 DB 682 TCCCTGACTACTGAGGCGCAAGGACACGCTCACCGTCTCTCCGAGGTGGATGCC 741  
 QY 98 ----- 98  
 DB 742 GAGGTGCAAGCTCTGAGTCTGAGAGTGGCTGTGTCAGACCTGAGAGATCCCTGAATCTC 801  
 QY 98 ----- 98  
 DB 802 TCCTGTGAGCCTCAGATTGATTTAGATAGTATGATGATGATGGGTCCGAGAGCT 861  
 QY 99 ----- 99  
 DB 862 CCAGGAAAGGCTTGAATGATGATGAGAAATTAATCAAGATAGACATGAATAACTAT 921  
 QY 99 ----- 99

DB 922 AGCCATCTTAAGATAGATTCATCTCCAGACAAAGCCAAAATACGCTGAC 981  
 QY 100 -----GluValIleHis----- 103  
 DB 982 CTGCAATAGCAAGATGAGGTCTGAGACACAGCCCTTTATTACTGTCAAGATTGGGG 1041  
 QY 104 -----GlyTrpValProTrpIleGlnIleThrThrValThrValSerSerGlyIle 121  
 DB 1042 CAATGGGGTACTTGTACTTGGGGCCAGAGGACACGCTGACGCTCTCTCAGTGGT 1101  
 QY 122 GlyIleSerGlyGlyGlySerGlyGlyGlySerAspIleGlnLeuThrGlnSer 141  
 DB 1102 GGTGGTCTGGCGCGCGGCTCCGGTGGTGGTCTTGAGCTCGTATACACAGTCT 1161  
 QY 142 HisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysValAspSerGln 161  
 DB 1162 CCATCTCTCTGACGTGACAGCAGAGAGGGGCACTATGAGCTGCAAGTCCAGTACG 1221  
 QY 162 AspValIleAspAla-----ValAlaTrpIleGlnIleLysProGly 175  
 DB 1222 AGTCTGTTAAACAGTGAATCAAAAGAACTACTGACCTGTACACAGAAACCAAGG 1281  
 QY 176 GlnSerProLysLeuLeuIleTrpSerAlaSerSerArgTrpThrGlyValProSerArg 195  
 DB 1282 CAGCTCTTAACCTGTGATCTACTGGGCACTCCATCGAAGATCGGGGTCTCATGTCG 1341  
 QY 196 PheThrGlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnIleGln 215  
 DB 1342 TTCACAGGCACTGATGAGAACAGATTCACTCTCACATCAGACGATGAGCTGATA 1401  
 QY 216 AspLeuAlaValIlePheCysGlnIleHisPheArgThrProPheThrPheGlySerGly 235  
 DB 1402 GACCTGGCAGTTATTACTGTGACAGATGATATATTATCTCTCACGCTTGGGTCTGG 1461  
 QY 236 ThrLysLeuGlnIleLys 241  
 DB 1462 ACCAAGCTTGAGATCAAA 1479

RESULT 13

US-09-766-543-9  
 ; Sequence 9, Application US/09766543  
 ; Patent No. US20020041865A1  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Austin, Richard  
 ; APPLICANT: Kwok, Cheuk S.  
 ; APPLICANT: Ring, David B.  
 ; TITLE OF INVENTION: METHODS FOR TREATING TUMORS  
 ; FILE REFERENCE: PP01679.002  
 ; CURRENT APPLICATION NUMBER: US/09/766,543  
 ; CURRENT FILING DATE: 2000-01-19  
 ; PRIOR APPLICATION NUMBER: 60/117,258  
 ; PRIOR FILING DATE: 2000-01-20  
 ; NUMBER OF SEQ ID NOS: 14  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 9  
 ; LENGTH: 816  
 ; TYPE: DNA  
 ; ORGANISM: Artificial Sequence  
 ; FEATURE:  
 ; OTHER INFORMATION: Description of Artificial Sequence: 520C9  
 ; OTHER INFORMATION: humanized single-chain antibody used in the  
 ; OTHER INFORMATION: IL-2-antibody fusions  
 ; NAME/KEY: CDS  
 ; LOCATION: (7) ..(807)  
 US-09-766-543-9

Alignment Scores:

Pred. No.: 7 036-86 Length: 816  
 Score: 913.00 Matches: 167  
 Percent Similarity: 83.61% Conservative: 37  
 Best Local Similarity: 68.44% Mismatches: 36  
 Query Match: 38.93% Indels: 4  
 DB: 9 Gaps: 2

US-09-596-774-7 (1-443) x US-09-766-543-9 (1-816)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuValProGlyGluThrValIysIle 20
Db 73 GAGATGCAACCTGGTGGAGCTGGGCTGAGGAAAGCCCTGGGCTCAGTGAAGCTC 132
QY 21 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIysGln 40
Db 133 TCCTGCAAGCTTCGCTTACACCTTACCACTTGAATGATCTGGGTGGCAAGCTC 192
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
Db 193 CCTGAGCAAGGCTTGAAGGATGGATGATGATTAACACTACACTGACAGTCAACATAT 252
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 253 GCTGATGACTTCAAGAAAGAGTCAACATGACACAGACATCCACGACGACGCTTAC 312
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db 313 ATGACACCTGAGAGAGCTGAGATCTACACAGACAGGCTGTATTAAGTCTGCGAGA----- 366
QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
Db 367 ---CGATTGGGATTGGCT---TACTGGGGCCAGGAGACCTGTGACCGCTCTCCCTCAGGT 420
QY 121 GlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlyGlySerAspIleGlnLeuThrGln 140
Db 421 GGGGTGGCTGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG 480
QY 141 SerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCysLysAlaSer 160
Db 481 TCTCATCTCTCCCTGCTGTCATCTGTGAGAGACAGAGTCACTACCTGCGGGGCAAGT 540
QY 161 GlnAspValTyrAsnAlaValAlaTyrTrpGlnGlnLysProGlyGlnSerProLysLeu 180
Db 541 CAGGACATGGTAATAGCTTAACCTGTATCAGACAGAAACAGGAGAAACCCCTAAGCTC 600
QY 181 LeuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThrGlySerGly 200
Db 601 CTGATCTACGCGCATCCAGCTTGAATTCGGGGTCCCATCAAGGTTCAAGGAAAGTGA 660
QY 201 SerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeuAlaValTyr 220
Db 661 TCTGGGCAAGATTTTACTTTCATCATCAGCAGTCTGCAAGCTGAAATATTGCAACATAT 720
QY 221 PheCysGlnGlnIleAspPheArgThrProPheThrPheGlySerGlyThrLysLeuGluIle 240
Db 721 TACTGCTTACAAATATGCTATTTTCCGTACACGTTCCGCCAAGGACACGACGAGATT 780
QY 241 LysAlaLeuGlu 244
Db 781 AAGGATCCGAA 792

```

RESULT 14  
US-10-127-890-90

Sequence 90, Application US/10127890  
Publication No. US20030166196A1

GENERAL INFORMATION:  
APPLICANT: Better, Marc D.  
Carroll, Stephen F.

Studilka, Gary M.

TITLE OF INVENTION: Immunosolins Comprising Ribosome-Inactivating  
Proteins

NUMBER OF SEQUENCES: 173  
CORRESPONDENCE ADDRESS:  
ADDRESSEE: McAndrews, Held & Malloy, Ltd.

STREET: 500 West Madison Street, 34th floor  
CITY: Chicago  
STATE: Illinois  
COUNTRY: USA  
ZIP: 60661

# COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: Patentin Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/10/127, 890

FILING DATE: 23-Apr-2002

CLASSIFICATION: <Unknown>

PRIOR APPLICATION DATA:

APPLICATION NUMBER: US/08/646, 360

FILING DATE: 13-MAY-1996

APPLICATION NUMBER: PCT/US94/05348

FILING DATE: 12-MAY-1994

APPLICATION NUMBER: US 08/064, 691

FILING DATE: 12-MAY-1993

APPLICATION NUMBER: US 07/984, 430

FILING DATE: 09-DEC-1992

APPLICATION NUMBER: US 07/901, 707

FILING DATE: 19-JUN-1992

APPLICATION NUMBER: US 07/787, 567

FILING DATE: 04-NOV-1991

ATTORNEY/AGENT INFORMATION:

NAME: McNicholas, Janet M.

REGISTRATION NUMBER: 32,918

REFERENCE/DOCKET NUMBER: 200-70.P4

TELEPHONE: 312/707-8889

TELEFAX: 312/707-9155

INFORMATION FOR SEQ ID NO: 90:

SEQUENCE CHARACTERISTICS:

LENGTH: 723 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

MOLECULE TYPE: DNA

SEQUENCE DESCRIPTION: SEQ ID NO: 90:

US-10-127-890-90

US-09-596-774-7 (1-443) x US-10-127-890-90 (1-723)

```

QY 1 GlnIleGlnLeuValGlnSerGlyProGluLeuValProGlyGluThrValIysIle 20
Db 1 GAGATCAGTGGTGGAGCTGGAGAGGCTGGTGAACCTGGAGGCTCGTCAAAATC 60
QY 21 SerCysAlaSerGlyTyrProPheThrAsnTyrGlyMetAsnTrpValIysGln 40
Db 61 TCTGGGCAAGCTTCGCTTACACCTTACCAACTATGATGATTAAGTCTGGTGGCAAGCT 120
QY 41 ProGlyGlnGlyLeuLysTrpMetGlyTyrIleAsnThrSerThrGlyGluSerThrPhe 60
Db 121 CCAGAAAGGTTTGAAGTGGATGGCTGATTAACACCACTGAGAGCCCAACATAT 180
QY 61 AlaAspAspPheLysGlyArgPheAspPheSerLeuGluThrSerAlaAsnThrAlaTyr 80
Db 181 GCTGATCTTCAAGGAGCGGTTTACCTTCTTTGGAGATTCCTTAAGAACCTGCTTAT 240
QY 81 LeuGlnIleAsnAsnLeuLysSerGlyLysAspMetAlaThrTyrPheCysAlaArgTrpGlu 100
Db 241 TTACAGATTAAGAGCTGAGACGACGACGACGACGACGACGACGACGACGACGACGAC 297
QY 101 ValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrValSerSerGly 120
Db 298 GGTACAGCTGCTACTGATGCTGGGGCCAAAGGAGCCACGCTCACCGTCTCTCAGGT 357

```



Thu Apr 21 06:48:36 2005

us-09-596-774-7.apr20.rnpb

Page 14

Db : 718 AAA 720

Search completed: April 20, 2005, 17:07:55  
Job time : 781.098 secs

---

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 20, 2005, 06:53:52 ; Search time 17.1515 Seconds  
(without alignments)  
2485.142 Million cell updates/sec

Title: US-09-596-774-7  
Perfect score: 2345  
Sequence: 1 QIOLVSGPELKKRGETVKI.....LSTATKDYDALHMQTLAPR 443

Scoring table: BIOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%

Listing first 45 summaries

Database : PIR 79:\*  
1: pir1:\*  
2: pir2:\*  
3: pir3:\*  
4: pir4:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	713	30.4	164	2	A40104 T-cell receptor CD
2	694	29.6	268	2	A56446 Ig heavy chain V r
3	665.5	28.4	249	2	S41374 single chain Fv an
4	611.5	26.1	233	2	JC5322 p53 specific singl
5	605	25.8	177	2	S54817 T-cell receptor io
6	605	25.8	206	2	A35900 T-cell receptor CD
7	604.5	25.8	163	2	A31768 T-cell receptor ze
8	589.5	25.1	188	2	A45089 T-cell receptor CD
9	539	23.0	166	2	JC4664 T-cell surface gly
10	537.5	22.9	120	2	B42848 L6 mab heavy chain
11	537.5	22.8	146	4	S32190 Ig heavy chain pre
12	535.5	22.8	146	4	S32190 Ig heavy chain pre
13	525	22.4	136	2	S35755 BHD9D10 protein -
14	522.5	22.3	118	2	S37204 Ig heavy chain V r
15	519	22.1	119	2	A53285 Ig heavy chain V a
16	516	22.0	119	2	H45722 anti-glycoprotein
17	510	21.7	117	2	S32190 Ig heavy chain V r
18	499.5	21.3	115	2	S19665 Ig heavy chain V r
19	495	21.1	107	2	S32192 Ig kappa chain V r
20	490	20.9	107	2	S32191 Ig kappa chain V r
21	490	20.9	113	2	B36259 Ig heavy chain V r
22	488.5	20.8	118	2	A32530 Ig heavy chain V r
23	488.5	20.8	139	2	PH1225 Ig heavy chain pre
24	487	20.8	124	2	PH1404 Ig heavy chain V r
25	486.5	20.7	115	2	S19968 Ig heavy chain V r
26	486	20.7	117	2	S32187 Ig heavy chain V r
27	486	20.7	118	2	S19967 Ig heavy chain V r
28	486	20.7	149	1	KVMS11 Ig kappa chain pre
29	484.5	20.7	109	2	S26325 Ig heavy chain V r

30	484.5	20.7	120	2	S19963 Ig heavy chain V r
31	484	20.6	119	2	B32530 Ig heavy chain V r
32	480	20.5	119	2	P00265 Ig kappa chain V r
33	477.5	20.4	114	2	D32967 Ig heavy chain V r
34	472.5	20.1	114	2	C32967 Ig heavy chain V r
35	469	20.0	152	2	S30751 Ig kappa chain pre
36	460	19.6	114	2	P00265 Ig heavy chain V r
37	453.5	19.3	108	2	PL02083 Ig kappa chain V r
38	452	19.3	117	2	S42466 Ig kappa chain V r
39	451	19.2	131	2	S26792 Ig heavy chain V r
40	450	19.2	105	2	S24765 Ig heavy chain V r
41	448.5	19.1	127	2	S04577 Ig kappa chain pre
42	446	19.0	131	2	PL0207 Ig heavy chain V r
43	443.5	18.9	101	2	D24672 Ig heavy chain V r
44	442	18.8	107	2	S09967 Ig kappa chain V-J
45	442	18.8	134	2	S21916 Ig heavy chain V r

## ALIGNMENTS

RESULT 1  
A40104  
T-cell receptor CD3 zeta chain - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 20-Mar-1992 #sequence\_revision 20-Mar-1992 #text\_change 09-Jul-2004  
A:Accession: A40104; I55293  
R:Weisman, A.M.; Baniyash, M.; Hou, D.; Samelson, L.E.; Burgess, W.H.; Klausner, R.D.  
A:Title: Molecular cloning of the zeta chain of the T cell antigen receptor.  
A:Reference number: A40104; MUID:88145643; PMID:3278377  
A:Accession: A40104  
A:Status: preliminary  
A:Molecule type: mRNA  
A:Residues: 1-164 <WEI>  
A:Cross-references: UNIPROT:P24161; GB:M19729; NID:G201131; PID:AAA40171.1; PID:G201132  
R:Baniyash, M.; Heu, V.W.; Seldin, M.F.; Klausner, R.D.  
J. Biol. Chem. 264, 13252-13257, 1989  
A:Title: The isolation and characterization of the murine T cell antigen receptor zeta cl  
A:Reference number: I55293; MUID:88327299; PMID:2787796  
A:Accession: I55293  
A:Status: preliminary; translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-164 <RES>  
A:Cross-references: GB:J04967; NID:G556326; PID:AAA50301.1; PID:G556327  
A:Gene: Tcrz  
C:Keywords: phosphoprotein; T-cell receptor; transmembrane protein  
Query Match 30.4%; Score 713; DB 2; Length 164;  
Matches 138; Similarity 99.3%; Pred. No. 5.4e-43; Indels 0; Gaps 0;  
Mismatch 1; Conserved 1;  
OY 305 LEDPKCYLIDGLIFIVGVIITLALYLRKFSRSRSETANLQDPNOLYNELNIGRREYDV 364  
DB 26 LIDPKCYLIDGLIFIVGVIITLALYLRKFSRSRSETANLQDPNOLYNELNIGRREYDV 85  
OY 365 LEKRRADPEWKGKQQRNRRPQEGVYNALQDKMAEYSEIGTGERRRGKHGGLYQGL 424  
DB 86 LEKRRADPEWKGKQQRNRRPQEGVYNALQDKMAEYSEIGTGERRRGKHGGLYQGL 145  
OY 425 STATKDYDALHMQTLAPR 443  
DB 146 STATKDYDALHMQTLAPR 164  
RESULT 2  
A56446  
Ig heavy chain V region (3H-3H scFv) - mouse (strain BALB/C)  
C:Species: Mus musculus (house mouse)  
C:Date: 19-Jan-1996 #sequence\_revision 19-Jan-1996 #text\_change 16-Aug-1996  
C:Accession: A56446  
R:Tang, P.M.; Foltz, L.A.; Mahoney, W.C.; Schueler, P.A.



C:\Species: Mus musculus (house mouse)  
C:\Date: 23-Oct-1990 #sequence\_revision 23-Oct-1990 #text\_change 09-Jul-2004  
C:\Accession: A35900; A60374; G46522  
R:\Jin, Y.-Y.; Clayton, L.K.; Howard, F.D.; Koyasu, S.; Sieh, M.; Steinbrich, R.; Tarr, G.  
Proc. Natl. Acad. Sci. U.S.A. 87, 3319-3323, 1990  
A:\Title: Molecular cloning of the CD3zeta subunit identifies a CD3zeta-related product in  
A:\Reference number: A35900; MUID:90239005; PMID:2139725  
A:\Accession: A35900  
A:\Status: preliminary  
A:\Molecule type: mRNA  
A:\Residues: 1-206 <JIN>  
A:\Cross-references: UNIPROT:P29020; GB:M33158; NID:g192488; PIDN:AAA37398.1; PID:g309155  
R:\Ono, H.; Saito, T.  
Int. Immunol. 2, 1117-1119, 1990  
A:\Title: CD3zeta and eta chains are produced by alternative splicing from a common gene.  
A:\Reference number: A60374; MUID:91190781; PMID:2150596  
A:\Accession: A60374  
A:\Status: not compared with conceptual translation  
A:\Molecule type: DNA  
A:\Residues: 144-206 <OHN>  
R:\Jensen, J.P.; Conclarielli, C.; Hou, D.; Relhanan, B.L.; Dean, M.; Weissman, A.M.  
J. Immunol. 150, 122-130, 1993  
A:\Title: T cell antigen receptor-eta subunit. Low levels of expression and limited cross  
A:\Reference number: A46522; MUID:93107707; PMID:8411118  
A:\Contents: annotation  
C:\Comment: The functional significance of this alternatively spliced product of the CD3  
ino acids, differs widely among various mammalian species in sequence, length, and even  
C:\Keywords: alternative splicing; T-cell receptor; transmembrane protein

Query Match            25.8%; Score 605; DB 2; Length 206;  
Best Local Similarity   99.2%; Pred. No. 2.6e-35;  
Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Oy         305 LEDEKLCYLDDGILFIYGVITLALYLRAKFSRSASETAANLDPPQLYNELNLGRREEDV 364  
Db         26 LLDPLCLTLDDGILFIYGVITLALYLRAKFSRSASETAANLDPPQLYNELNLGRREEDV 85

Oy         365 LEKKRADPEWGGKQOORRRNPQEGVYNALQDKMAEAYSEIGTGERRRGKHDLTYQ 422  
Db         86 LEKKRADPEWGGKQOORRRNPQEGVYNALQDKMAEAYSEIGTGERRRGKHDLTYQ 143

RESULT 7  
A31768      T-cell receptor zeta chain precursor - human  
C:\Species: Homo sapiens (man)  
C:\Date: 07-Jun-1990 #sequence\_revision 07-Jun-1990 #text\_change 09-Jul-2004  
C:\Accession: A31768  
R:\Weissman, A.M.; Hou, D.; Orloff, D.G.; Modi, W.S.; Seunarez, H.; O'Brien, S.J.; Klausne  
Proc. Natl. Acad. Sci. U.S.A. 85, 9709-9713, 1988  
A:\Title: Molecular cloning and chromosomal localization of the human T-cell receptor zet  
A:\Reference number: A31768; MUID:89071765; PMID:2974162  
A:\Accession: A31768  
A:\Molecule type: mRNA  
A:\Residues: 1-163 <WEI>  
A:\Cross-references: UNIPROT:P20963; GB:J04135; NID:9623041; PIDN:AAA60394.1; PID:9623042  
C:\Keywords: phosphoprotein; T-cell receptor; transmembrane protein  
F:\1-21\Domain: signal sequence #status predicted <SIG>  
F:\2-163/Product: T-cell receptor zeta chain #status predicted <MAT>

Query Match            25.8%; Score 604.5; DB 2; Length 163;  
Best Local Similarity   84.9%; Pred. No. 2.1e-35;  
Matches 118; Conservative 6; Mismatches 14; Indels 1; Gaps 1;

Oy         305 LEDPKLCYLDDGILFIYGVITLALYLRAKFSRSASETAANLDPPQLYNELNLGRREEDV 364  
Db         26 LLDPLCLTLDDGILFIYGVITLALFLRKFSRSABPPYQGQOLYNELNLGRREEDV 85

Oy         365 LEKKRADPEWGGKQOORRRNPQEGVYNALQDKMAEAYSEIGTGERRRGKHDLTYQGL 424  
Db         86 LDKRGRDPEWGGK-PRIKNQEBGLYNELQDKMAEAYSEIGTMGERRRGKHDLTYQGL 144

Oy         425 STATKDYDALHMQTAPR 443

Db 145 STATKDYDALHMQALPDR 163

RESULT 8

A45089  
T-cell receptor CD3 theta chain, alternate splice form - mouse  
C:Species: Mus musculus (house mouse)  
C:Date: 30-Apr-1993 #sequence\_revision 18-Nov-1994 #ext\_change 05-Nov-1999  
C:Accession: A45089; 149587  
R:Clayton, L.K.; Diener, A.C.; Lerner, A.; Tse, A.G.; Koyasu, S.; Reinherz, E.L.  
J. Biol. Chem. 267, 26023-26030, 1992  
A:Title: Differential regulation of T-cell receptor processing and surface expression aff  
A:Reference number: A45089; MUID:93100325; PMID:146413  
A:Accession: A45089  
A:Status: Preliminary; not compared with conceptual translation  
A:Molecule type: mRNA  
A:Residues: 1-188 <CLA>  
A:Cross-references: GB:S51932; NID:g261998; PIDN:AAB24559.1; PID:g261999  
A:Experimental source: thymus  
A>Note: Sequence extracted from NCBI backbone (NCBI:P120865)  
R:Lerner, A.; Diener, A.C.; Reinherz, E.L.; Clayton, L.K.  
Eur. J. Immunol. 22, 2135-2140, 1992  
A:Title: Human genomic sequences corresponding to murine CD3eta-related transcripts: lack  
A:Reference number: 149587; MUID:92347411; PMID:1322304  
A:Accession: 149587  
A:Status: translated from GB/EMBL/DBJ  
A:Molecule type: mRNA  
A:Residues: 1-188 <RES>  
A:Cross-references: GB:L03353; NID:g192508; PIDN:AAA37401.1; PID:g192509  
C:Keywords: T-cell receptor

Query Match 25.1%; Score 589.5; DB 2; Length 188;  
Best Local Similarity 98.3%; Pred. No. 2,8e-34;  
Matches 116; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

Oy 305 LEDPRLCYLLDGLGILFYGVITLALYLRAKFSRSAAETPAALQDPNOLYNELNGRRREYDV 364  
Db 26 LDDPRLCYLLDGLGILFYGVITLALYLRAKFSRSAAETPAALQDPNOLYNELNGRRREYDV 85

Oy 365 LEKRRARPENMGKQKQRRNPOEGVYNALQKQKMAEAYSEIGTKGRRRKGHDGLYQ 422  
Db 86 LEKRRARPENMGKQKQRRNPOEGVYNALQKQKMAEAYSEIGTKGRRRKGHDGLYQ 142

RESULT 9

JC4664  
T-cell receptor zeta chain - bovine  
C:Species: Bos primigenius taurus (cattle)  
C:Date: 10-May-1996 #sequence\_revision 19-Jul-1996 #ext\_change 05-Nov-1999  
C:Accession: JC4664  
R:Hagens, G.; Galley, Y.; Glaser, I.; Davis, W.C.; Baldwin, C.L.; Clevers, H.; Dobbelaere  
Gene 169, 165-171, 1996  
A:Title: Cloning, sequencing and expression of the bovine CD3 epsilon and TCR-zeta chainm  
A:Reference number: JC4663; MUID:96194796; PMID:8647441  
A:Accession: JC4664  
A:Molecule type: mRNA  
A:Residues: 1-166 <HAG>  
C:Cross-references: GB:U5668; NID:g1263011; PIDN:AAC48548.1; PID:g1263012  
C:Comment: This protein plays a pivotal role in linking T-cell receptor-triggering to se  
ymphokine receptor gene expression.  
C:Genetics:  
A:Gene: tcr-zeta  
C:Keywords: GTP binding; signal transduction; T-cell receptor  
F:129-146/Region: GDP/GTP-binding

Query Match 23.0%; Score 539; DB 2; Length 166;  
Best Local Similarity 75.2%; Pred. No. 8,3e-31;  
Matches 106; Conservative 11; Mismatches 22; Indels 2; Gaps 1;

Oy 305 LEDPRLCYLLDGLGILFYGVITLALYLRAKFSRSAAETPAALQDPNOLYNELNGRRREYDV 364  
Db 26 LDDPRLCYLLDGLGILFYGVITLALYLRAKFSRSAAETPAALQDPNOLYNELNGRRREYDV 85

```

QY      365 LEKRRAPENMGKGKQORRNPOEGVYNLQDKMAEAYSEIGTK--GERRRGKHGDGLYO 422
       :::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
Db      86 LDRGGPPEPGSGKGRQRKKGNPREVVYINELRCKDAEAYSSEIGMSKDNRRRGKHGDGLYO 145
       |||||
QY      423 GLSTATKDTYDALHMOTLA PR 443
       |||||
Db      146 GLSTATKDTYDALHMQALPR 166
       |||||

RESULT 10
I46424
T-cell surface glycoprotein CD3 zeta chain - sheep
C:Species: Ovis orientalis aries, Ovis ammon aries (domestic sheep)
C>Date: 16-Aug-1996 #sequence_revision 16-Aug-1996 #text_change 09-Jul-2004
C:Accession: I46424; S22980
R:Hein, W.R.; Tunnaciffie, A.
Immunogenetics 37, 279-284, 1993
A>Title: Invariant components of the sheep T-cell antigen receptor: cloning of the CD3 e
A:Reference number: I46424; MUID:93131305; PMID:8420837
A:Accession: I46424
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-166 <HEI>
A:Cross-references: UNIPROT:P29329; EMBL:Z12968; NID:g1399; PIDN:CAA78312.1; PID:g1400
Keywords: glycoprotein
```

Query Match	23.0%	Score 539	DB 2	Length 166
Best Local Similarity	74.5%	Pred. No. 8.3e-31		
Matches 105	Conservative 13	Mismatches 21	Indels 2	Gaps 1

QY 305 LEEPKLCYLLDGLIFLYGYIITATYLRKAFKFSASATAANIQDPNOLVNEINLGRREEDY 364  
 Db 26 LDDPKCYLLDGLIFLYGYIITATFLRAKFSASADAPAYQCGQNPVYNELNVRREYAV 85  
 QY 365 LKKKARDBEMGSKOORPNEPQEGYNNLODKMAEAYSETGT--GERRGKCHDLQY 422  
 Db 86 LDRRGSGFDEMGSKQORKKNPEHYVNNELRDKMAEAYSETGKSDNORRGKCHDGYQ 145  
 QY 423 GLSTATKOTYDALHMQTLAPR 443  
 Db 146 GLSTATKOTYDALHMQALPRL 166

```

RESULT 11
B42848
L6 mAb heavy chain V region - mouse (fragment)
C/Species: Mus musculus (house mouse)
C/Date: 27-Apr-1993 #sequence_revision 18-Nov-1994 #text_change 21-Jan-2000
A/Accession: B42848, S33503
R/Fell, H.P.; Gayle, M.A.; Yelton, D.; Lipsich, L.; Schieven, G.L.; Marken, J.S.; Aruffo
J. Biol. Chem. 267, 15552-15558, 1992
A/Title: Chimeric L6 anti-tumor antibody. Genomic construction, expression, and character
A/Reference number: A42848, MUID:92348410, PMID:1633794
A/Accession: B42848
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-120 <FE1>
A/Cross-references: GB:M90690; NID:G195065; PIDN:AAA38146.1; PID:G195066
A/Note: sequence extracted from NCBI backbone (NCBIN:109960, NCBIPI:109961)
A/Accession: S33903
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-120 <FE2>
A/Cross-references: EMBL:M90691
C/Superfamily: Immunoglobulin V region; immunoglobulin homology <IMM>
P:15-98/Domain: immunoglobulin homology <IMM>

```

Query Match	22.9%;	Score 537.5;	DB 2;	Length 120;
Best Local Similarity	84.2%;	Pred. No. 7.1e-31;		
Matches 101; Conservative	7;	Mismatches 11;	Indels 1;	Gaps 1;

QY 1 QIQLVQSGPELKKPGETVKISCKASGYPFTNYGMNVKQAPGQGLKWMGWINTSTGESTF 60

Db 1 QIDLVSGBELKKKPGHTVAKISCKASGYITPTTNGNMVYQACGKGLKMGWINTTGGPTY 60

Qy 61 ADDFKGRFDPFSLSTANTAYIQINNLKSHDVAIFYCAEWEYHG-VYPYWGQGTTLTVSS 119

Db 61 ADDFKGRFAPSLSTSTAYTIQINNLKSHDVAITYFCARFSYGNSRIVADYWGQGTTLTVSS 120

RESULT 12  
S33905  
Ig heavy chain precursor V region - synthetic  
Accession Number

C/Date: 13-Jan-1995 #sequence\_revision 30-Apr-1998 #ext\_change 20-Oct-2000  
C/Accession: S33905  
R/LIU, A.Y.; Robinson, R.R.; Hellstrom, K.E.; Murray Jr., E.D.; Chang, C.P.; Hellstrom  
Proc. Natl. Acad. Sci. U.S.A. 84, 3439-3443, 1987  
A/Title: Chimeric mouse-human IgG1 antibody that can mediate lysis of cancer cells.  
A/Reference number: S33905; WCID:87204152; PMID:3106570  
A/Accession: S33905  
A/Molecule type: mRNA  
A/Residues: 1-146 <LIU>  
A/Cross-references: EMBL:M16072; NID:g195270; PIDD:AAA98229.1; PID:g195271

Query Match	22.8%;	Score 535.5;	DB 4;	Length 146;
Best Local Similarity	84.2%;	Pred. No. 1.2e-30;		
Matches 101; Conservative	7;	Mismatches 11;	Indels 1;	Gaps 1

[illegible]

**QY** 61 ADDPKGRDPSLETSANTAYLQINNLSKSEMATYFCARWEYYHG-VYPYWGOGTTVWSS 119  
||||| : ||||| : ||||| : |||||  
**Db** 80 ADDEFGRAFSLETSAYTAYLQINNLKNEDMATYFCARFSGNSRSDYWGOGTTLTWSS 139

```

RESULT 13
S35759
EBHD9D10 protein - mouse (fragment)
C:Species: Mus musculus (house mouse)
C:Date: 13-Jan-1995 #sequence_revision 13-Jan-1995 #text_change 21-Jan-2000
C:Accession: S35759
R:Froeyn, G.F.V.
submitted to the EMBL Data Library, May 1993
A:Reference number: S35759
A:Accession: S35759
A:Status: preliminary
A:Residue type: mRNA
A:Molecule: 1-136 <PRO>
A:Cross-references: EMBL:X72796; NID:g312496; PIDN:CAA51336.1; PID:g312497
E:34-117/Domain: immunoglobulin homology <1MM>
C:Superfamily: immunoglobulin V region; immunoglobulin homology

```

Query Match	22.4%	Score 525;	DB 2;	Length 136;
Best Local Similarity	84.0%	Pred. No. 6.2e-30;		
Matches 100;	Conservative	6;	Mismatches 11;	Indels 2;
				Gaps 1

QY QIOLVSGPELKKRGETVTKISKASGYPLFNNGMMNVKAPGGGLKWMGMINTSGEGRF 60  
Db 20 QIOLVSGPELKKRGETVTKISKASGYPLFDIGMMNVKAPGGGLKWMGMINTSGEGRF 79  
QY 61 ADFFKGRFDESLTSANTAYLQINLNKSEDMATYFCARVEVHGVPYWGQGTVTSS 119  
80 VDFEKKRFAFSLTSSAAAYLQINLNKMDKDTYTPCARGFV-ANDYWGQGTSTVSS 136

RESULT 14  
S37204

Ig heavy chain V region  
C/Species: Mus musculus (house mouse)  
C/Date: 19-Mar-1997 #sequence\_revision 19-Mar-1997 #text\_change 20-Jun-2000  
C/Accession: S37204  
C/Fischer: R.; Voss, A.; Hunziker, W.; Stierhof, Y.D.; Kreuzaler, F.



submitted to the EMBL Data Library, August 1993

A;Reference number: S37200

A: Accession: S37204

A; STATUS: preliminary  
A; ACCEPTED: 05/20/

A: molecule type: mRNA

A;Molecule type: mRNA

A;Residues: 1-118 <FIS>

A/Cross-references: EMBL:X74589; NID:g402599; PIDN:CAA52666.1; PID:g1333981

C: Superfamily: immunoglobulin V region; immunoglobulin homology

F:15-98/Domain: immunoglobulin homology <IMM>

Query Match	Score	DB 2;	Length
22.34;	522.5;	DB 2;	Length 118;
80.74;	Pred No	7 98-30;	

Best Local Similarity 80.7%; Pred. No. 7.8e-30;

Matches 96; Conservative 9; Mismatches 13; Indels 1; Gaps 1;

Dy  
1 QIQLVQSGPELKKPGETVYKISCSASGYPTNTYGMMWVKAPGCGLKMGWINTSTGESETF 600

Dz  
1 QIQLVQSGPELKKPGETVYKISCSASGYFTIDYSMHWKQVPGCKGLKMGMWINTETGEAKY 600

---

61 ADDEGKRFARSLETSASTAYLLQNNLKKEDTATYFCARWG-YSANFPYMGCGTLVTVSA 118

### RESULT 15

A53285

Ig heavy chain V and J regions, monoclonal antibody SCET.M8.1 - mouse (fragment)

C:\Species: Mus musculus (house mouse)

```
C:\species: was musculus (mouse mouse)
C:\Date: 02-May-1994 #sequence revision 18-Nov-1994 #text change 20-Jun-2000
```

C; Date: 02-May-1994  
C; Accession: A53285

C/Accession: A53285

R; Sawada, J.; Mizubawa, S.; Terao, T.; Naito, M.; Kurosawa, Y.

Mol. Immunol. 28, 1063-1072, 1991

A; Title: Molecular characterization

and their pH-reactivity profiles.

A;Reference number: 1

A: Accession: A53285

A;ACCESSION: A53285  
A;STATUS: preliminary

A; Status: preliminary

A; molecule type: DNA; protein

A;Residues: 1-119 <SAW>

A; Cross-references: GB:D12736; NID:g220595; PIDN:BAA02228.1; PI

A;Note: sequence extracted from NCBI backb

C:Superfamily: immunoglobulin V region; immun

**Keywords:** heterotetramer; immunoglobulin F<sub>15</sub>-98/Domain: immunoglobulin homology <IMM>

[illegible]

Query Match 22.18; Score 519; DB 2; Length 119;

Matches	98; Conservative	8; Mismatches	13; Indels	0; Gaps
QY	1 QIQLVSGPELKKPGETVKISCSASGYPPTNYGMWVKAPGGGLKMMGWINTSTGSETF	60		

[illegible]

**Qy**      61    ADDEKGRFDPSLETSANTAYLOJNLKSEDMATYFCARBEVHYGVYPWGGOTTVYSS    119  
| | | | | : | | | | | : | | | | | : | | | | |  
**Pb**      61    ADDEKGRFAESLETSTASTAYLOJNDLKNEEDATYFCARGNVRYVANDYWGOSTSVYSS    119  
| | | | | : | | | | | : | | | | | : | | | | |

01 ADDFNKRFHSEI SAS IHLQINDENED IHI IFCAGN VKV IHIWIMWYISV I VSS I I I

Search completed: April 20, 2005, 07:07:49  
Job time : 18.1515 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame\_plus\_p2n model

Run on: April 20, 2005, 10:13:45 ; Search time 4346.69 Seconds  
(without alignments)  
3879.380 Million cell updates/sec

Title: US-09-596-774-7  
Perfect score: 2345  
Sequence: 1 QIOLVOSGPEIKKPGFTVKI.....LSTATKDYALHMOTIAPR 443

Scoring table:  
BLOSUM62  
Xgapop 10.0 , Xgapext 0.5  
Ygapop 10.0 , Ygapext 0.5  
Fgapop 6.0 , Fgapext 7.0  
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues  
Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Command line parameters:

-MODE=frame+;p2n.model; DBY=xlh  
-Q=/cg2\_1/USPTO.spool/h/US09596774/runac.20042005.075443.12521/app.query.fasta\_1.1230  
-DB=EST -QPM=fastap -SUFF1=apr20.trst -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0  
-UNITS=bits -START=1 -END=1 -MATRIX=blousum62 -TRANS=human40.cdi -LIST=45  
-DOCALLGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL  
-OUTPM=pic -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000  
-USRR=US09596774.@CGN 1.1\_4591.@runac.20042005.075443.12521 -NCPU=6 -ICPU=3  
-NO\_MMAP -LARGESUBSTRY -NEG\_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG  
-DEV TIMEOUT=120 -MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6  
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:\*  
1: gb\_eest1.\*  
2: gb\_eest2.\*  
3: gb\_hc.\*  
4: gb\_eest3.\*  
5: gb\_eest4.\*  
6: gb\_eest5.\*  
7: gb\_eest6.\*  
8: gb\_gsest1.\*  
9: gb\_gsest2.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	746.5	31.8	671	3	AF240168 Mus muscu
2	715	30.5	953	6	BY746048 BY746048
3	715	30.5	1570	3	AK088087 Mus muscu
4	713	30.4	577	6	CA577711 K0713C05
5	704.5	30.0	1598	3	AK017904 Mus muscu
6	697.5	29.7	492	9	AT403844 Mus muscu
7	681.5	29.1	764	7	C0568334 AGENCOURT
8	681.5	29.1	786	7	C0572391 AGENCOURT
9	628	26.8	993	4	BI832527 603082128

c	10	621	26.5	872	7	CK629396	CK629396 AMO-AA001
	11	620	26.4	618	7	CV029513	CV029513 8329 Full1
	12	620	26.4	772	4	BI838213	BI838213 603083131
	13	620	26.4	1038	5	BM921412	BM921412 AGENCOURT
	14	620	26.4	1113	1	AL554350	AL554350 AL554350
	15	620	26.4	1546	3	CR625818	CR625818 full-1eng
	16	620	26.4	1557	3	CR601423	CR601423 full-1eng
	17	617.5	26.3	547	5	B0562923	B0562923 H4078G04-
	18	615	26.2	587	5	BP367042	BP367042 BP367042
	19	614	26.2	1025	1	AL558350	AL558350 AL558350
	20	607.5	25.9	899	7	CO64852	CO64852 ILLUMIGEN
	21	606.5	25.9	601	5	B0474958	B0474958 carabue4g
	22	604.5	25.8	492	9	AY403842	AY403842 Homo sapi
	23	604.5	25.8	575	5	BP367299	BP367299 BP367299
	24	604.5	25.8	581	5	BP368604	BP368604 BP368604
	25	604.5	25.8	712	1	AL557555	AL557555 AL557555
	26	604.5	25.8	875	5	BX464301	BX464301 BX464301
	27	604.5	25.8	970	5	B0051777	B0051777 AGENCOURT
	28	604.5	25.8	1095	1	AL532484	AL532484 AL532484
	29	604.5	25.8	1126	3	CR615998	CR615998 full-1eng
	30	602.5	25.7	1104	7	CK629846	CK629846 AM2-AA002
c	31	596.5	25.4	580	5	BP367951	BP367951 BP367951
	32	578.5	24.7	581	5	BP366659	BP366659 BP366659
	33	567.5	24.2	852	4	BI819337	BI819337 603035487
	34	567	24.2	672	7	CK633068	CK633068 AM3-AP001
	35	567	24.2	672	7	CK633069	CK633069 AM3-AP001
	36	565	24.1	675	5	BP163731	BP163731 BP163731
	37	565	24.1	707	5	BP167815	BP167815 BP167815
	38	565	24.1	727	5	BP164018	BP164018 BP164018
	39	565	24.1	823	5	BP167326	BP167326 BP167326
	40	565	24.1	835	5	BP163661	BP163661 BP163661
	41	565	24.1	841	5	BP161017	BP161017 BP161017
	42	565	24.1	841	5	BP164042	BP164042 BP164042
	43	565	24.1	845	5	BP434480	BP434480 BP434480
	44	565	24.1	885	5	BP161782	BP161782 BP161782
	45	564	24.1	695	5	BP166127	BP166127 BP166127

## ALIGNMENTS

RESULT 1  
AF240168  
LOCUS AF240168 671 bp mRNA linear HTC 30-APR-2001  
DEFINITION Mus musculus MRP5 mRNA, partial cds.  
ACCESSION AF240168  
VERSION AF240168.1 GI:13877286  
KEYWORDS  
SOURCE  
ORGANISM  
Mus musculus (house mouse)  
HTC.  
Mus musculus

REFERENCE  
AUTHORS  
Cui,D.X., Zeng,G.Y., Wang,F., Xu,J.R., Ren,D.O., Guo,Y.H.,  
Tian,F.R., Yan,X.J., Hou,Y., Ren,D.O., Guo,Y.H.,  
Mechanism of exogenous nucleic acids and their precursors improving  
the repair of intestinal epithelium after gamma-irradiation in mice  
World J. Gastroenterol. 6 (5), 709-717 (2000)

JOURNAL  
PUBMED  
11819679  
2 (bases 1 to 671)  
Cui,D., Zeng,G., Yan,X., Li,X. and Su,C.  
Cloning of mouse genes related to repairing of intestinal  
epithelium of the irradiated mice by treatment with the intestinal  
RNA of mice of the same strain  
Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19 (1), 71-80  
(2001)

REFERENCE  
AUTHORS  
Cui,D., Zeng,G., Yan,X., Wang,F., Tian,F., Ren,D., Zhao,T., Li,X.  
and Su,C.  
Direct Submission  
Submitted (02-MAR-2000) Molecular Cloning Laboratory, Institute of  
Genetic Diagnosis, Fourth Military Medical University, No. 17 Chang  
Le West Road, Xi'an 710032, China

JOURNAL  
TITLE  
JOURNAL  
Le West Road, Xi'an 710032, China  
Location/Qualifiers

FEATURES  
3 (bases 1 to 671)  
Cui,D., Zeng,G., Yan,X., Wang,F., Tian,F., Ren,D., Zhao,T., Li,X.  
and Su,C.  
Direct Submission  
Submitted (02-MAR-2000) Molecular Cloning Laboratory, Institute of  
Genetic Diagnosis, Fourth Military Medical University, No. 17 Chang  
Le West Road, Xi'an 710032, China  
Location/Qualifiers

source  
 1. 671.  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="BALB/c"  
 /db\_xref="taxon:10090"  
 /sex="male"  
 /cell\_type="intestinal gland cell"  
 /note="derived from differential display PCR between irradiated mice with and without RNA injection"  
 18. >671  
 /note="related to the repair of irradiation-damaged intestinal gland cells treated with RNA"  
 /codon\_start=1  
 /product="MRP5"  
 /protein\_id="AAK43733.1"  
 /db\_xref="GI:138772894"  
 /translation="MAQYKLOQSGELKKPGETVIRISKAGCYFTTAGMOWQPMGKGLKWTGINTHSVPTKAEERKFAFSLSTAYLQISLNKEDTATYFCMRMDYDGFAYWGQGTIVTSSGGSGSGSDIVLTQSPASIVASVIGSRATISCSA SESVDNIGISFWMWFQKPGQPRLIIYAASKSGSGVPAGLLASGSDTFSLNTPME"

## ORIGIN

Alignment Scores:  
 Pred. No.: 3,996-64 Length: 671  
 Score: 746.50 Matches: 141  
 Percent Similarity: 79.34% Conservative: 28  
 Best Local Similarity: 66.20% Mismatches: 39  
 Query Match: 31.83% Indels: 5  
 Gaps: 2

US-09-596-774-7 (1-443) x AF240168 (1-671)

QY 1 GlnileglnleuValGlnSerglYProgluLeuylsYProglYgluThrValylle 20  
 |||:|||||:  
 Db 24 CAGGTGAACGCGACGACTCAGACCTGACTGAAGAGCTGAGACAGCTAGATC 83  
 QY 21 SerCysylsAlaserglYTyPProPheThrAnTYrglMetAsnTPValysGlnAla 40  
 |||:|||||:  
 Db 84 TCCTGCAAGGCTTCGATATACCTTCACACTGCTGATGATGATGATGATGATG 143  
 QY 41 ProglYnglYleuYserTmeGlyTrpIleAnThrSerThrglYgluSerThPhe 60  
 |||:|||||:  
 Db 144 CAGGAAAGGATTTGAAGTGAATGGCTGATTAACACCACTGATGATGCCAAAGT 203  
 QY 61 AlaAspPhePheYglYArgPheAspPheSerLeuGluThrSerAlaAsnThAlaTYr 80  
 |||:|||||:  
 Db 204 GCAGAGAGCTTCAGAGGACGCTTTCCTTCCTTGGAAAGCTTCGACAGCTGATAT 263  
 QY 81 LeuGlnIleAsnAnleuYSerGluAspMetAlaThrTYrPheCYsAlaArgTrpGlu 100  
 |||:|||||:  
 Db 264 TTACAGATTAAGCACTCAAAATAGAGACACGCTGATTTCTGTATGATGATGAT 323  
 QY 101 ValYThrHsGlyTYrValProTYrTrpGlyGlnIlyThrThrValIThrValSerSergly 120  
 |||:|||||:  
 Db 324 ---TACAGCGGGGGCTTCTTACTGCGGGCCAAAGGACACAGTCCCTCTCAGGT 380  
 QY 121 GlyYglYserGlyYglYglYserGlyYglYglYserAspIleGlnleuThGln 140  
 |||:|||||:  
 Db 381 GGAAGCGGCTTCAGCGGAGGCTGCGGAGTGGCGGATCGGACATCGGCTCACCAG 440  
 QY 141 SerIlsYsPheleuSerThrSerValGlyAspArgValSerIleThrCYsYsAlaSer 160  
 |||:|||||:  
 Db 441 TCTTCACGCTTCTTGGCTGTCTCTCTAGGGCAGAGGGCCACATCTCTCAGAGCAGC 500  
 QY 161 GlnAspValTYrAsnAla-----ValAlaThrTYrGlnIlyAspProglYglu 176  
 |||:|||||:  
 Db 501 GAAAGTGTATATATATGCGATTAGTTATGACTGTTCAGAGAAACAGAGAGAG 560  
 QY 177 SerProYsleuLeuIleTYrSerAlaSerSerArgTYrThrglyValProSerArgPhe 196  
 |||:|||||:  
 Db 561 CCACCCAAAGCTCTATCTATGCTGATCCAGACAGAGATCCGGGGTCTCTGACAGTTTA 620

QY 197 ThrGlySerGlySerGlyProAspPheThrPheThrIle 209  
 |||:|||||:  
 Db 621 CTGGCAAGTGGGCTCTGGACAGATTTTCAGCTTCAACATA 659

RESULT 2  
 BY746048  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM

REFERENCE  
 AUTHORS

197 ThrGlySerGlySerGlyProAspPheThrPheThrIle 209  
 621 CTGGCAAGTGGGCTCTGGACAGATTTTCAGCTTCAACATA 659

BY746048 953 bp mRNA linear EST 17-DEC-2002  
 BY746048 RIKEN full-length enriched, 2 days neonate thymus thymic  
 cells (NOD) Mus musculus cDNA clone B430003J03 5', mRNA sequence.  
 BY746048.1 GI:27173533  
 EST.  
 Mus musculus (house mouse)

Eukaryote; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 1 (bases 1 to 953)  
 Okazaki, Y., Furuno, M., Kasukawa, T., Adachi, J., Bono, H., Kondo, S.,  
 Nishikado, I., Osato, N., Saito, R., Suzuki, H., Yamanaka, I.,  
 Kiyosawa, H., Yagi, K., Tomaru, Y., Hasegawa, Y., Nogami, A.,  
 Schombach, C., Gojohori, T., Balcarrelli, R., Hill, D. P., Bull, C.,  
 Hume, D. A., Quackenbush, J., Schrym, L. M., Kanapin, A., Matsuda, H.,  
 Batalov, S., Beisel, K. W., Blake, J. A., Brad, D., Brusic, V.,  
 Chothia, C., Corbani, L. E., Cousins, S., Dalla, E., Dragan, T. A.,  
 Fletcher, C. F., Forrest, A., Frazer, K. S., Gaasterland, T.,  
 Gariboldi, M., Gissi, C., Godzik, A., Gough, J., Grimmond, S.,  
 Gustinchik, S., Hirokawa, N., Jackson, I. J., Jarvie, E. D., Kanai, A.,  
 Kawasaki, H., Kawasawa, Y., Kedziarski, R. M., King, B. L., Konagaya, A.,  
 Kurochkin, I. V., Lee, Y., Lennard, B., Lyons, P. A., Maglott, D. R.,  
 Maltais, L., Marchionni, L., McKenzie, L., Miki, H., Nagashima, T.,  
 Numata, K., Okido, T., Pavan, W. J., Perlea, G., Pesole, G.,  
 Petrovsky, N., Pillai, R., Pontius, J. U., Qi, D., Ramachandran, S.,  
 Ravasi, T., Reed, J. C., Reed, D. J., Reid, J., Ring, B. Z., Ringwald, M.,  
 Sandelin, A., Schneider, C., Semp, C. A., Seton, M., Shinada, K.,  
 Siltana, R., Takenaka, Y., Taylor, M. S., Teasdale, R. D., Tomita, K.,  
 Verardo, R., Wagner, L., Wahlestedt, C., Wang, Y., Watanabe, Y.,  
 Wells, C., Wilming, L. G., Wymshaw-Boris, A., Yanagisawa, M., Yang, I.,  
 Yang, L., Yuan, Z., Zavolan, M., Zhu, Y., Zimmer, A., Carninci, P.,  
 Hayatsu, N., Hirozane-Kishikawa, T., Kono, H., Nakamura, M.,  
 Sakazume, N., Sato, K., Shiraki, T., Waki, K., Kawai, J., Aizawa, K.,  
 Arakawa, T., Fukuda, S., Hara, A., Hashizume, M., Imotani, K., Ishii, Y.,  
 Itoh, M., Kagawa, I., Miyazaki, A., Sakai, K., Sasaki, T., Shibata, K.,  
 Shinagawa, A., Yasunishi, A., Yoshino, M., Waterston, R., Lander, E. S.,  
 Rogers, J., Birney, E. and Hayashizaki, Y.  
 Analysis of the mouse transcriptome based on functional annotation  
 of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)

JOURNAL  
 MEDLINE  
 PUBMED

12466851  
 12466851

COMMENT  
 Contact: Yoshihide Hayashizaki  
 Laboratory for Genome Exploration Research Group, RIKEN Genomic  
 Sciences Center (GSC), Yokohama Institute  
 The Institute of Physical and Chemical Research (RIKEN)  
 1-7-22 Suenho-cho, Tsunumi-ku, Yokohama, Kanagawa 230-0045, Japan  
 Tel: 81-45-503-9222  
 Fax: 81-45-503-9216  
 Email: genome-res@gs.c.riken.jp, URL: http://genome.gsc.riken.jp/  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Carninci, P.,  
 Fukuda, S., Hashizume, W., Hayashida, K., Hirozane, T., Hori, F.,  
 Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kawai, J., Kojima, Y.,  
 Kondo, S., Kono, H., Koya, S., Miyazaki, A., Murata, M., Nakamura, M.,  
 Nomura, K., Numazaki, R., Ohno, M., Osato, N., Saito, R., Sakazume, N.,  
 Sano, H., Sasaki, D., Sato, K., Shibata, K., Shiraki, T., Tagami, M.,  
 Takeda, Y., Waki, K., Watanabe, A., Muramatsu, M. and Hayashizaki, Y.  
 Direct Submision  
 Computational Analysis of Full-length Mouse cDNAs Compared with  
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)  
 Normalization and subtraction of cap-trapper-selected cDNAs to  
 prepare full-length cDNA libraries for rapid discovery of new  
 genes. Genome Res. 10 (10), 1617-1630 (2000)  
 RIKEN integrated sequence analysis (RISA) system--384-format  
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.  
 10 (11), 1757-1771 (2000)

Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)  
 A CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.  
 Tissues were provided by Dr. John Todd (Dept. of Medical Genetics Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome Trust/MRC building Addenbrookes Hospital Cambridge) whose assistance we gratefully acknowledge.  
 Please visit our web site (<http://genome.gsc.riken.go.jp>) for further details.

# FEATURES

location/Qualifiers

1..953  
 /organism="Mus musculus"  
 /mol\_type="mRNA"  
 /strain="NOD"  
 /db\_xref="taxon:10090"  
 /clone="E43003L03"  
 /tissue\_type="thymus"  
 /cell\_type="thymic cells"  
 /clone\_id="RIKEN full-length enriched, 2 days neonate thymus thymic cells (NOD)"

# ORIGIN

## Alignment Scores:

Pred. No.: 9.36e-61 Length: 953  
 Score: 715.00 Matches: 146  
 Percent Similarity: 86.39% Conservative: 0  
 Best Local Similarity: 86.39% Mismatches: 11  
 Query Match: 30.49% Indels: 12  
 Gaps: 2

US-09-596-774-7 (1-443) x BY746048 (1-953)

275 SerProValHisProThrGlyThrsSerGlnProGlnArgProGluAapCyAspProGlu 294  
 94 TCGCTGCTGATCTCCACCTGCGGCTCCAGAGCAG-----AGCAGCAGGA 138  
 295 GlySerValIyGlyGlyThrGlyLeuAspPheLeuGluAapProGlyLeuCyAspTyLeuLeu 314  
 139 GCTTTGGCTCG-----CTGATCCCAACTCTGCTACTTGGTA 177  
 315 AspGlyIleLeuPheIleTyGlyValIleIleThrAlaLeuTyLeuArgAlaIyPhe 334  
 178 GATGGAATCCCTCTCATCTACGAGATCATCAAGCCCTGTACTGAGAGCAAAATTC 237  
 335 SerArgSerIaGluThrAlaAlaAsnLeuGlnAapProAsnGlnLeuTyAsnGluLeu 354  
 238 AGCAGAGTGAGAGAGACTGCTGCCAAGGAGGAGCCCAACAGCTTACAAATGAGCTC 297  
 355 AsnLeuGlyArgArgGluGluTyArgValLeuGlyIyAspValAArgAspProGlu 374  
 298 AATTAAGGCGCAAGAAGAGATATGACTCTTGAAGAAGAGCGGCTCGGATCCAGAG 357  
 375 MetGlyIyGlyGlnGlnArgArgAspProGlnGluGlyValTyAsnAlaLeuGln 394  
 358 ATGGAGAGCAAAACGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 417  
 395 LysAspIyAspMetAlaGluAlaTySerGlnIyGlyIyThrIyGlyIyGluArgArgGly 414  
 418 AATAACAAGATGGAG 477  
 415 LysGlyIyHisAspGlyLeuTyGlnGlyLeuSerThrAlaThrIyAspThrTyAspAla 434  
 478 AAGGAGGACATGAGCTTTACAGAGGTCTCAGCACTGCCACCAAGAGAGAGAGAGAGAG 537  
 435 LeuHisMetGlnThrLeuAlaProArg 443  
 538 CTGCATATGACAGACCTGGCCCTCGC 564

## RESULT 3

AK088087 1570 bp mRNA linear HTC 03-APR-2004  
 LOCUS AK088087

## DEFINITION

Mus musculus 2 days neonate thymus thymic cells cDNA, RIKEN full-length enriched library, clone:E43003L03 product:CD3 antigen, zeta polypeptide, full insert sequence.

## ACCESSION

AK088087.1 GI:26104673

## VERSION

HTC: CAP trapper.

## KEYWORDS

Mus musculus (house mouse)

## SOURCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

## REFERENCE

1 Carninci, P. and Hayashizaki, Y.  
 High-efficiency full-length cDNA cloning  
 Meth. Enzymol. 303, 19-44 (1999)

## JOURNAL

Genome Res. 10 (10), 1617-1630 (2000)

## PUBMED

20499374

## TITLE

3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komuro, H., Akiyama, J., Nishi, K., Kikunishi, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsunoto, H., Sakaguchi, S., Ikegami, T., Kaishiwagi, K., Fujiwara, S., Inoue, K., Togawa, Y., Iwawa, M., Ohara, E., Matsubiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A., and Hayashizaki, Y.  
 RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer  
 Genome Res. 10 (11), 1757-1771 (2000)

## JOURNAL

Genome Res. 10 (11), 1757-1771 (2000)

## PUBMED

11042159

## TITLE

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.  
 Functional annotation of a full-length mouse cDNA collection  
 Nature 409, 685-690 (2001)

## JOURNAL

Nature 409, 685-690 (2001)

## PUBMED

11076861

## TITLE

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.  
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs  
 Nature 420, 563-573 (2002)

## JOURNAL

Nature 420, 563-573 (2002)

## PUBMED

12000000

## TITLE

6 (Bases 1 to 1570)  
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirose, T., Hirose, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kigawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komuro, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numasaki, R., Ohno, M., Ohnishi, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y.

## JOURNAL

Submitted (16-APR-2002) Yoshinori Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-gsc@riken.jp, URL: <http://genome.gsc.riken.jp/>, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

## COMMENT

CDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken  
Genomic Sciences Center and Genome Science Laboratory in RIKEN.  
Division of Experimental Animal Research in Riken contributed to  
prepare mouse tissues.  
Tissues were provided by Dr. John Todd (Dept. of Medical Genetics  
Wellcome Trust Centre for Molecular Mechanisms in Disease Wellcome  
Trust/MRC building Addenbrookes Hospital Cambridge) whose  
assistance we gratefully acknowledge.  
Please visit our web site for further details.  
URL: <http://genome.gsc.riken.jp/>.  
URL: <http://phantom.gsc.riken.jp/>.  
Location/Qualifiers

## FEATURES

source

1..1570  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="NOD"  
/db\_xref="FANTOM DB:E430003103"  
/db\_xref="taxon:10090"  
/clone="E430003103"  
/cell\_type="thymic cells"  
/tissue\_type="thymus"  
/clone\_id="RIKEN full-length enriched mouse cDNA library"  
/dev\_stage="2 days neonate"  
76..571  
/note="CD3 antigen, zeta polypeptide (MGD|MG1:88334,  
GB|NM\_031162, evidence: BLASTN, 100%, match=536)  
putative"  
1549..1554  
/note="putative"  
1570  
/note="putative"

## ORIGIN

## Alignment Scores:

Pred. No.: 1..9e-60 Length: 1570  
Percent Similarity: 715.00 Matches: 146  
Best Local Similarity: 86.39% Conserv: 0  
Query Match: 30.49% Mismatches: 11  
DB: 3 Gaps: 2

US-09-596-774-7 (1-443) x AK088087 (1-1570)

275 SerProValHisProthnGlyThrSerGlnProGlnArgProGluAspCysArgProArg 294  
98 TCGGCTGATCCTCCACGTCGCGTTCCAGAGAGC-----AGGCACAGA 142  
295 GlySerVallysglyThnrglyLeuaspPheleuGluaspProlysluGlyCysTyrLeuLeu 314  
143 GCTTGTGCTCG-----CTGATGCCAAACTCTGCTACTCTCTA 181  
315 AspGlylleuPheleuPheleuVallelleuThraaleuTyrleuArgalaValPhe 334  
182 GATGATCCCTTCATCTACGAGATCATCTACAGCCCTTGACGAGACAAATTC 241  
335 SerArgSerAlaGluThrAlaAlaAsnleuGlnaaspProanglnleuTyrAsnGluLeu 354  
242 AGCAGAGTGCAGAGACGTGCTGCCAAGCTGCAGAGCCCAACGCTCTCAATGAGCTC 301  
355 AsnleuGlyArgArgGluGluTyrAspValleuGluTyrslsArgAlaAaaspProGlu 374  
302 AATCTAAGGCAAGAGAGAAATGACCTCTGGAAGAAAGCGGCTCGGATCCAG 361  
375 MetGlyGlyslnglnArgArgAspProGlnGluGlyValTyrAsnAlaLeuGln 394  
362 ATGAGAGCAAAACGACAGAGAGAGAACCCCAAGAGCGGTATCAATGCACTGAG 421  
395 LysAspLysMetAlaGluAlaTyrSerGlnIleGlyThrLysGlyGluArgArgGly 414  
422 AAAGACAAGATGGCAGAGCCTACAGATGAGATCGCAAAAGCGGAGCGAGAGGC 481  
415 LysGlyHisAspGlyLeuTyrGlnGlyLeuSerThrAlaThrLysAspThrTyrAspAla 434

Db 482 AAGGGACACAGTGCCTTTACAGGCTCTGACGACCTGCCACAGACCACTATGATGCC 541

Oy 435 LeuHisMetGlnThrLeuAlaProArg 443

Db 542 CTGCATATGACGACCTTGCCCTCGC 568

## RESULT 4

## LOCUS

CA577711 577 bp mRNA linear EST 19-NOV-2002

K0713C05-5N NIA Mouse Hematopoietic Stem Cell (Lin-/c-Kit-/Sca-1-)

IMAGE:30074812 5', mRNA sequence.

## ACCESSION

CA577711.1 GI:25126102

## VERSION

## KEYWORDS

## SOURCE

## ORGANISM

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

Other\_ESTs: K0713C05-3  
Contact: Dawood B. Dudekula  
Laboratory of Genetics  
National Institute on Aging/National Institutes of Health  
333 Cassell Drive, Suite 4000, Baltimore, MD 21224-6820, USA  
Email: cdna@gsun.gsc.nia.nih.gov  
Plate: K0713 row: C column: 05  
Seq primer: M13 Reverse  
High quality sequence stop: 577  
POLYA=No.

## FEATURES

## source

Location/Qualifiers  
1..577  
/organism="Mus musculus"  
/mol\_type="mRNA"  
/strain="C57BL/6NCr"  
/db\_xref="nlaEST:K0713C05-5N"  
/db\_xref="taxon:10090"  
/clone="NIA:K0713C05 IMAGE:30074812"  
/tissue\_type="Hematopoietic Stem Cell  
(Lin-/c-Kit-/Sca-1-)"  
/dev\_stage="Age approx. 10 weeks old"  
/lab\_host="DH10B"  
/clone\_id="NIA Mouse Hematopoietic Stem Cell  
(Lin-/c-Kit-/Sca-1-)"  
/note="Vector: pSPORT (Invitrogen); Site 1: SalI; Site 2:  
NotI; Mouse cDNA project by the Laboratory of Genetics,  
National Institute on Aging (NIA), Intramural Research  
Program, NIH (<http://19sun.gsc.nia.nih.gov/cDNA>). This is  
a long-transcript enriched cDNA library (Ref. Genome Res.  
11: 1553-1558 (2001). [PMID: 11544191]). Total RNAs were  
obtained from Drs. Dennis Taub, Dan Longo (National  
Institute on Aging, USA), Jonathan Keller (National Cancer  
Institute, USA). Double-stranded cDNAs were synthesized  
with an Oligo(dT) primer (Invitrogen:  
5'-pACTAGTCTTAGATCGCAGCGCCCTTTT-3') from  
0.9 ug of total RNA, treated with T4 DNA polymerase, and  
purified by ethanol-precipitation. The cDNAs were ligated  
to lone-linker IL-Sal4, purified by phenol/chloroform, and  
separated from free linkers by Centricon 100. Then, the  
cDNAs were amplified by long-range high fidelity PCR using  
ex Tag polymerase (Takara) with a primer Sal4-S. The  
products were purified by phenol/chloroform and Centricon  
100. The cDNAs were digested with SalI and NotI enzymes  
and cloned into SalI/NotI site of pSPORT plasmid vector.  
The DH10B E. coli host was transformed with the ligation  
mixture by the standard chemical method. The average  
insert size is about 2.1 Kb. The library was constructed  
by Yulan Piao (NIA)."

TITLE	Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuzaki, S., Kawai, J., Ozaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y.
JOURNAL	RIKEN integrated sequence analysis (RISA) system--384-format
MEDLINE	sequencing pipeline with 384 multicapillary sequencer
PUBMED	Genome Res. 10 (11), 1757-1771 (2000)
REFERENCE	20530913
AUTHORS	11076861
TITLE	4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium.
JOURNAL	Functional annotation of a full-length mouse cDNA collection
REFERENCE	Nature 409, 685-690 (2001)
AUTHORS	5
TITLE	The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.
JOURNAL	Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
REFERENCE	Nature 420, 563-573 (2002)
AUTHORS	6 (bases 1 to 1598)
TITLE	Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Fukunishi, Y., Furuno, M., Hanagaki, T., Hara, A., Hayatsu, N., Hiramoto, K., Hirakawa, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Izawa, M., Kikukawa, T., Kato, H., Kawai, J., Kojima, Y., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Okido, T., Owa, C., Saito, H., Saito, R., Sakai, C., Sakai, K., Sano, H., Saeki, D., Shibata, K., Shibata, Y., Shingawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, Y., Tagawa, A., Takahashi, F., Tanaka, T., Tajima, Y., Toyota, T., Yamamura, T., Yasunishi, A., Yoshida, K., Yoshino, M., Muramatsu, M. and Hayashizaki, Y.
TITLE	Direct Submission
JOURNAL	Submitted (10-JUL-2000) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)
COMMENT	Please visit our web site (http://genome.gsc.riken.jp/) for further details.
TITLE	cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. First strand cDNA was primed with a primer [5', GAGAGAGACGATCCGAGTAAATTAATTAATTCACCCCCC 3']. cDNA was prepared by using trehalase thermo-activated reverse transcriptase and subsequently enriched for full-length by cap-trapper. cDNA went through one round of normalization to Rot = 10.0 and subtraction to Rot = 100.0. Second strand cDNA was prepared with the primer adapter of sequence [5', GAGAGAGACATTCGAGTAAATTAATTAATTCACCCCCC 3']. cDNA was cleaved with BamHI and XhoI. Vector: a modified plasmid KS(+) after bulk excision from lambda FLC I. Cloning sites, 5' end: SalI, 3' end: BamHI. Host: DH10B.
FEATURES	Location/Qualifiers
SOURCE	1. 1598
FEATURES	/organism="Mus musculus"
SOURCE	/mol_type="mRNA"
FEATURES	/strain="C57BL/6J"
SOURCE	/db_xref="PANTOM_DB:5830407D18"
FEATURES	/db_xref="taxon:10090"
SOURCE	/clone="5830407D18"
FEATURES	/sex="male"
SOURCE	/tissue_type="thymus"
FEATURES	/clone_lib="RIKEN full-length enriched mouse cDNA library"
SOURCE	/dev_stage="adult"
FEATURES	112. 606
SOURCE	/note="unnamed protein product: CD3 antigen, zeta polypeptide (MGDI MG1:88334, GB NM_031162, evidence: BLASTN, 100%, match=536)
SOURCE	putative"





IMAGE:7384006 5', mRNA sequence.  
 ACCESSION COS68334  
 VERSION COS68334.1 GI:50380963  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 764)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics / NIH  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgsapbs-remail.nih.gov  
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical  
 College of Wisconsin  
 CDNA Library Preparation: Open Biosystems  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLM15548 row: d column: 20  
 High quality sequence stop: 746.  
 Location/Qualifiers

FEATURES  
 source  
 1..764  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="IMAGE:7384006"  
 /lab\_host="DH10B Tona"  
 /clone\_lib="NIH MGC 250"  
 /note="Organ: thymus; Vector: pExpress-1; Site 1: EcoRV;  
 Site 2: NotI; RNA obtained from testis tissue of 8 wk old  
 animal. Tissues were snap-frozen and kept at -80C before  
 RNA extraction and purification (TRI-reagent method). CDNA  
 5'-pGACTAGTCTAGATCGGAGCGGCCGC(T)25-3' and cloned into  
 the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb  
 resulted in an average insert size of 1.9 kb. This is a  
 primary library (normalized library is NIH MGC 251) and  
 was constructed by Open Biosystems. Note: this is a  
 NIH\_MGC library"

# ORIGIN

Alignment Scores: 1.55e-57 Length: 764  
 Pred. NO.: 681.50 Matches: 140  
 Score: 81.56% Conservative: 6  
 Percent Similarity: 78.21% Mismatches: 21  
 Best Local Similarity: 29.06% Indels: 12  
 Query Match: 7 Gaps: 3

US-09-596-774-7 (1-443) x COS68334 (1-764)  
 Oy 274 ProserProvalHisPro-----ThrgLYThrSerLInPro 285  
 Db 28 CCGCTCTGGGACACCTCCAGGCTTCAGAAAGTGAAGCGAGCATCAGCTCGC 87  
 Oy 286 GlnAArgProGluAArgCyAArgProArgGlySerValLySGlyThrgLYleu-AArgPheLe 305  
 Db 88 CTGCATCTCTAAGTACGATTCGCCAGAGC-----AGAGCGACAGACCTTGGCTCTCT 141  
 Oy 305 uGluAArgProLyseuCyetyrLeuLeuAArgGlylleuPheleTyrgLYValIle1 325  
 Db 142 G---GATCCCAAACTCTCTATATGCTATGATGAAATCTCTTCACTACGAGATGATGT 198  
 Oy 325 eThrAlaLeuTyrlLeuAArgAlaLyPheSerArgSerAlaGluThrAlaAlaAsnLeuG1 345

Db 199 CACGGCCCTGTACTGAGACGAAATTACAGAGAGTGCAGATGCTGCTTACTTCA 258  
 Oy 345 nAsPProAArgnLIneutyrrAArgnGluLeuAAsnLeuGlyAArgGluGluTyrrAArgPvalLe 365  
 Db 259 GGACCCCAACGAGCTCTTATACGAGCTCAACTTGGCGGAGAGAGAAATGATGTTT 318  
 Oy 365 uGluLyLeuAArgAlaAArgAArgProGluMetGlyGlyLyseGlnGlnAArgAArgAArgPr 385  
 Db 319 GGACAGAGAGCGGCTCGGAGTCCGAGATGCGGGGCGAAACAGCAGAGAGAGAGAACCC 378  
 Oy 385 oGlnGlnGlyValTyrrAArgAlaLeuGlnLyAArgLyMetAlaGluAArgTyrrSerGlu1 405  
 Db 379 CCAGAAAGGCGCTGTACATGATGCTGCGAAAGACAAATGCGAGAGGCTTCACTGAGAT 438  
 Oy 405 eGlyThrgLYleuGlyGluAArgAArgGlyLySGlyVHisAArgPlyLeuTyrgLYnGlyLeu 425  
 Db 439 TGGCATGAAGAGCGAGAGCGGAGAGCGGAGGCGGAGCGGCTTTTACAGGATTCAG 498  
 Oy 425 rThrAlaThrgLYsAArgPThrgLYrAArgAlaLeuHisMetGlnThrgLYleuAArgAArg 443  
 Db 499 CACTGCCACACAGACACCTATGACGCCCTGCAATGCAAGCCCTGCCCTCGC 553

RESULT 8  
 AGENCOURT 28621196 NIH MGC 250 Rattus norvegicus CDNA clone  
 LOCUS IMAGE:7381111 5', mRNA sequence.  
 ACCESSION COS72391  
 VERSION COS72391.1 GI:50385020  
 KEYWORDS EST.  
 SOURCE Rattus norvegicus (Norway rat)  
 ORGANISM Rattus norvegicus  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;  
 Rattus.

REFERENCE 1 (bases 1 to 786)  
 NIH-MGC http://mgi.nci.nih.gov/.  
 National Institutes of Health, Mammalian Gene Collection (MGC)  
 Unpublished (1999)  
 CONTACT: Daniela S. Gerhard, Ph.D.  
 Office of Cancer Genomics  
 National Cancer Institute / NIH  
 Bldg. 31 Rm10A07 Bethesda, MD 20892  
 Email: cgsapbs-remail.nih.gov  
 Tissue Procurement: Drs. Josef Lazar & Howard Jacob, Medical  
 College of Wisconsin  
 CDNA Library Preparation: Open Biosystems  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Agencourt Bioscience Corporation  
 Clone distribution: MGC clone distribution information can be  
 found through the I.M.A.G.E. Consortium/LNL at:  
 http://image.llnl.gov  
 Plate: LLM15540 row: 1 column: 05  
 High quality sequence stop: 701.  
 Location/Qualifiers

1..786  
 /organism="Rattus norvegicus"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:10116"  
 /clone="IMAGE:7381111"  
 /lab\_host="DH10B Tona"  
 /clone\_lib="NIH MGC 250"  
 /note="Organ: thymus; Vector: pExpress-1; Site 1: EcoRV;  
 Site 2: NotI; RNA obtained from testis tissue of 8 wk old  
 animal. Tissues were snap-frozen and kept at -80C before  
 RNA extraction and purification (TRI-reagent method). CDNA  
 5'-pGACTAGTCTAGATCGGAGCGGCCGC(T)25-3' and cloned into  
 the EcoRV/NotI sites of pExpress-1. Size-selection >1.25kb  
 resulted in an average insert size of 1.9 kb. This is a  
 primary library (normalized library is NIH MGC 251) and  
 was constructed by Open Biosystems. Note: this is a  
 NIH\_MGC library"



VERSION CK629396.1 GI:45753871  
KEYWORDS EST.  
SOURCE Apis mellifera (honey bee)  
ORGANISM Apis mellifera  
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;  
Neoptera; Endopterygota; Hymenoptera; Apocrita; Aculeata; Apoidea;  
Apoidea; Apis.  
REFERENCE 1 (bases 1 to 872)  
Nunes, F.M.F., Valente, V., Sousa, J.F., Cunha, M.A.V., Pinheiro, D.G.,  
Mota, R.M., Araújo, D.D., Costa, M.C.R., Martins, W.K., Carvalho, A.F.,  
Meneses, N., Nascimento, A.M., Peixoto, P.M.V., Silva, M.F.R.,  
Ramos, R.G.P., Reis, L.F.L., Dias Neto, E., Souza, S.J., Simpson, A.J.G.,  
Zago, M.A., Soares, A.E.E., Bitondi, M.G., Espinola, E.M.,  
Espinola, F.S., Peco-Larson, M.L., Simoes, Z.L.P., Hartfelder, K. and  
Silva, M.A. Jr.  
TITLE Open reading frame ESTs - an efficient strategy for analysis of the  
honey bee transcriptome  
JOURNAL Unpublished (2004)  
COMMENT Contact: Silva Jr, W. A.  
Molecular Genetic and Bioinformatics Laboratory  
Department of Genetics, FMRP/USP, FUNDHERP  
Rua Tenente Celso Roxo, 2501, CEP 14051-140, Brazil  
Tel: +55 16 39639300  
Fax: +55 16 39639309  
Email: wilsonj@usp.br  
This sequence was derived from the FAPESP Genome Program  
High quality sequence start: 66  
Location/Qualifiers  
1. .872  
/organism="Apis mellifera"  
/mol\_type="mRNA"  
/strain="Africanized"  
/db\_xref="taxon:7460"  
/sex="female, worker"  
/dev\_stage="adult"  
/clone\_id="AA0013"  
/note="Organ: whole body"

ORIGIN  
Alignment Scores:  
Pred. No.: 2.15e-51 Length: 872  
Score: 621.00 Matches: 122  
Percent Similarity: 71.43% Conservative: 38  
Best Local Similarity: 54.46% Mismatches: 63  
Query Match: 26.48% Indels: 4  
DB: 7 Gaps: 1

US-09-596-774-7 (1-443) x CK629396 (1-872)

QY 18 ValysIleSerCylValAlaSerGlyTyrProPheThrAntyrGlyMetAsnTrpVal 37  
Db 775 ATGAAATGCTCTCCACAGCT-TCTGC-TACATCTTCACAGATGATATGATAGTACGCTG 718  
QY 38 LysGlnAlaPProGlyGlnGlyLeuLeuTyrPheMetGlyTTPileAsnThrSerThGlyGlu 57  
Db 717 AGGCAGACGCTGAACACGAGCTTGAAGTGGATGTTT-CTGAGAGAGGAGAGT 659  
QY 58 SerThrPheAlaAspPhePheLeuGlyTyrPheAspPheSerLeuGluThrSerAlaAsn 77  
Db 658 ACTGAATACATGAGAGAGTTCAGAGGCGAGCCACACATGATGAGACATCTCCGCGC 599  
QY 78 ThrAlaTyrLeuGlnIleAsnAsnLeuLeuSerGlyAspMetAlaThrTyrPheCysAla 97  
Db 598 ACACCTATATGAGGCTACTAGCTGACATCTAGAGCTGCTGCTATATTTCTGCT 539  
QY 98 ArgTrpGluValTyrHisGlyTyrValProTyrTrpGlyGlnGlyThrValThrVal 117  
Db 538 AGAGGAGCATCTATAGGCGCTACTTTGACTTTGGGNCACAGGACACCGCTACCGTC 479  
QY 118 SerSerGlyGlyGlyGlySerGlyGlyGlyGlySerGlyGlyGlySerAspIleGln 137  
Db 478 TCTCATGTGTGAGNCGGTTGAGCGGAGGTGGCTCTGGCGGTGGCGGATCTGACATTGAG 419

QY 138 LeuThrGlnSerHisLysPheLeuSerThrSerValGlyAspArgValSerIleThrCys 157  
Db 418 CTCACCCAGTCTCCAGCAATATCTGCATCTCCAGGAGAGAGGTCACCATGACCTGC 359  
QY 158 LysAlaSerGlnAspValTyrAsnAlaValAlaTyrGlnGlnIlyProGlyGlnSer 177  
Db 358 AGTCCAGCTCAAGTATA---CGTTACATATATGGTATCCACAGAACCCGATCTCC 302  
QY 178 ProLysLeuIleuIleTyrSerAlaSerSerArgTyrThrGlyValProSerArgPheThr 197  
Db 301 CCCAGACTCCGATTATGACACATCCAGCTGCTCTGAGTCCCTTTTTCGCTTCAGT 242  
QY 198 GlySerGlySerGlyProAspPheThrPheThrIleSerSerValGlnAlaGluAspLeu 217  
Db 241 GGCAGTGGTCTGGAGCTTATCTCTCATCATCAACGAATGAGAGGCTGAGATGCT 182  
QY 218 AlaValTyrPheCysGlnGlnHisPheArgThrProPheThrPheGlySerGlyThrLys 237  
Db 181 GCCACTTATTACTGCCAGAGTGGAGTGTATCTGTACACGTTCCGAGAGGCGGACCAAG 122  
QY 238 LeuGluIleLys 241  
Db 121 CTGAGGCTGAAA 110

RESULT 11  
CV029513 618 bp mRNA linear EST 20-AUG-2004  
LOCUS sapiens cDNA 5' similar to BC025703, mRNA sequence.  
DEFINITION  
ACCESSION CV029513  
VERSION CV029513.1 GI:51487684  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 618)  
Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S.,  
Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O.,  
Clingingsmith, T.R., Hartley, J.L., Espósito, D., Cheo, D., Moore, T.,  
Simmons, B., Sequerra, R., Bosak, S., Doucette-Stamm, L., Le Peuch, C.,  
Vanderhaute, J., Cusick, M.E., Albaladejo, J.S., Hill, D.E. and Vidal, M.  
Human ORFome Version 1.1: a Platform for Reverse Proteomics  
Genome Res. (2004) In press  
TITLE JOURNAL  
COMMENT Marc Vidal Laboratory  
Dana Farber Cancer Institute  
1 Jimmy Fund Way Smith 858, BOSTON, MA 02115, USA  
Tel: 617 632 5180  
Fax: 617 632 5739  
Email: Marc.Vidal@dfci.harvard.edu  
ORF Sequence Tag (OST) of Gateway Entry construct. Each cloned ORF  
results from a PCR reaction using an MGC full-length cDNA as  
template DNA and ORF specific primers  
PCR primers  
FORWARD: ATGAGTGAAGGCGCTT  
BACKWARD: TAGCAGGGGCGAGGCGCTG  
Insert Length: 618 Std Error: 29.00  
Plate: 11027 row: 09 column: H  
Seq primer: ACTGGCGCTGTTTACACAGCTGCTGACTGAGGAAAC  
High quality sequence start: 100  
PolyA=No.  
FEATURES  
Location/Qualifiers  
1. .618  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/tissue="mixed"  
/clone\_id="Full Length cDNA from the Mammalian Gene  
Collection"  
/note="Vector: mixed; The ORFs were PCR amplified from the

MGC (Mammalian Gene Collection) as of April 2004 and cloned by recombinational Gateway cloning into pDONR223 donor vector. Reference: MGC (Mammalian Gene Collection) Program Team, Generation and Initial Analysis of more than 15,000 Full-Length Human and Mouse cDNA Sequences. PNAS, 2002, 99(26), 16899-16903"

## ORIGIN

## Alignment Scores:

Pred. No.:	1,66e-51	Length:	618
Score:	620.00	Matches:	126
Percent Similarity:	79.64%	Conservative:	7
Best Local Similarity:	75.45%	Mismatches:	21
Query Match:	26.44%	Indels:	13
DB:	7	Gaps:	2

US-09-596-774-7 (1-443) x CVO29513 (1-618)

QY 278 HisProthrglyThrSerGlnProGlnArgProGluAspCyAspProArgGlySerVal 297  
 DB 30 CATCTCGAGGACAGCTTGGCGATTACAG-----59

QY 298 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuCySerGlyLeuAspGlyI 317  
 DB 60 ---GGCAGAGAGCTTGGCTGCTG---GATCCCAAACTCTGCTAAGCTGATGAGAT 113

QY 317 eleuPheIleTyrglyValIleIleThrAlaLeuTyreuArgAlaLysPheSerArgse 337  
 DB 114 CCTCTTCATCTAGTGTGTCATCTCCTGCTGAGAGTGAAGTTCACAGAG 173

QY 337 rAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyraAsnGluLeuAsnLeuG 357  
 DB 174 CGCAGACCCCGGCGTACAGAGGCGCAGACACAGCTATATACAGAGCTCAATCAGG 233

QY 357 YArGArgGluGluTyraAspValLeuGluLysLysArgAlaArgAspProGluMetGlyG 377  
 DB 234 ACGAAGAGAGAGATCGATGTTGTTGACAGACGTCGCGGAGACCTTGATGGGGG 293

QY 377 YLysGlnGlnArgArgAspProGlnGluGlyValTyraAsnAlaLeuGlnLysAspLy 397  
 DB 294 AAACCCGAGAGAGAGAAACCTTCAGAGAGCTGTACATATACCTGCAAGAAATTA 353

QY 397 sMetAlaGluAlaTyraSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyH 417  
 DB 354 GATGCGGAGGCTTACAGATGATGGATGAAGGCGGCGGAGGCGCAAGGGCA 413

QY 417 sAspGlyLeuTyraGlnGlyLeuSerThrAlaThrLysAspThrTyraAspAlaLeuHsme 437  
 DB 414 CGATGGCCTTTACAGAGGCTTCAGTACAGCCACAGGACCTTACAT 473

QY 437 tGlnThrLeuAlaProArg 443  
 DB 474 GCAGGCCCTGCCCCCTCGC 492

## RESULT 12

BI838213 772 bp mRNA linear EST 04-OCT-2001

LOCUS 60308313.F1 NIH\_MGC\_120 Homo sapiens cDNA clone IMAGE:5222574 5',

DEFINITION mRNA sequence.

ACCESSION BI838213

VERSION BI838213.1 GI:15949763

KEYWORDS EST.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens (human)

REFERENCE 1 (bases 1 to 772)

AUTHORS NIH-MGC http://mgi.nci.nih.gov/.

TITLE National Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT Contact: Robert Strausberg, Ph.D.

Email: cgabs-remail.nih.gov

Tissue Procurement: Life Technologies, Inc.

## FEATURES

## source

CDNA Library Preparation: Life Technologies, Inc.  
 CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)  
 DNA Sequencing by: Incyte Genomics, Inc.  
 Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LNL at:  
<http://image.lnl.gov>  
 Plate: LLM1559 row: m column: 07  
 High quality sequence stop: 767.  
 Location/Qualifiers

## 1. 772

/organism="Homo sapiens"

/mol\_type="mRNA"

/db\_xref="taxon:9606"

/clone="IMAGE:5222574"

/lab\_host="DH10B"

/clone\_lib="NIH\_MGC\_120"

/note="Organ: pooled pancreas and spleen; Vector: PCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA source anonymous pool of spleen and pancreas from 28 yo male. Library is oligo-dT primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-2.5 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 025. Note: this is a NIH\_MGC Library."

## ORIGIN

## Alignment Scores:

Pred. No.:	2.28e-51	Length:	772
Score:	620.00	Matches:	126
Percent Similarity:	79.64%	Conservative:	7
Best Local Similarity:	75.45%	Mismatches:	21
Query Match:	26.44%	Indels:	13
DB:	4	Gaps:	2

US-09-596-774-7 (1-443) x BI838213 (1-772)

QY 278 HisProthrglyThrSerGlnProGlnArgProGluAspCyAspProArgGlySerVal 297  
 DB 90 CATCTCGAGGACAGCTTGGCGATTACAG-----119

QY 298 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuCySerGlyLeuAspGlyI 317  
 DB 120 ---GGCAGAGAGCTTGGCTGCTG---GATCCAAACTCTGCTAAGCTGATGAGAT 173

QY 317 eleuPheIleTyrglyValIleIleThrAlaLeuTyreuArgAlaLysPheSerArgse 337  
 DB 174 CCTCTTCATCTAGTGTGTCATCTCCTGCTGAGAGTGAAGTTCAGAGGAG 233

QY 337 rAlaGluThrAlaAlaAsnLeuGlnAspProAsnGlnLeuTyraAsnGluLeuAsnLeuG 357  
 DB 234 ACGAAGAGAGAGATCGATGTTGTTGACAGACGTCGCGGAGACCTTGATGGGGG 293

QY 357 YArGArgGluGluTyraAspValLeuGluLysLysArgAlaArgAspProGluMetGlyG 377  
 DB 294 AAACCCGAGAGAGAGAAACCTTCAGAGAGCTGTACATATACCTGCAAGAAATTA 353

QY 377 YLysGlnGlnArgArgAspProGlnGluGlyValTyraAsnAlaLeuGlnLysAspLy 397  
 DB 354 GATGCGGAGGCTTACAGATGATGGATGAAGGCGGCGGAGGCGCAAGGGCA 413

QY 397 sMetAlaGluAlaTyraSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyH 417  
 DB 414 CGATGGCCTTTACAGAGGCTTCAGTACAGCCACAGGACCTTACAT 473

QY 417 sAspGlyLeuTyraGlnGlyLeuSerThrAlaThrLysAspThrTyraAspAlaLeuHsme 437  
 DB 474 CGATGGCCTTTACAGAGGCTTCAGTACAGCCACAGGACCTTACAT 533

QY 437 tGlnThrLeuAlaProArg 443  
 DB 534 GCAGGCCCTGCCCCCTCGC 552

RESULT 13  
LOCUS BM921412 1038 bp mRNA linear EST 12-MAR-2002  
DEFINITION AGNCOURT\_6626432 NIH\_MGC\_115 Homo sapiens cDNA clone IMAGE:5752844  
5', mRNA sequence.  
ACCESSION BM921412  
VERSION BM921412.1 GI:19371791  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1038)  
AUTHORS NIH-MGC http://mgi.nci.nih.gov/  
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)  
JOURNAL Unpublished (1999)  
COMMENT Contact: Robert Strausberg, Ph.D.  
Email: cgsapbs-remail.nih.gov  
Tissue Procurement: Life Technologies, Inc.  
CDNA Library Preparation: Life Technologies, Inc.  
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (ULNL)  
DNA Sequencing by: Agencourt Bioscience Corporation  
Clone distribution: MGC clone distribution information can be  
found through the I.M.A.G.E. Consortium/ULNL at:  
http://image.llnl.gov  
Plate: L1AM12787 row: k column: 21  
High quality sequence stop: 730.  
Location/Qualifiers  
1. 1038  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="IMAGE:5752844"  
/lab\_note="DH10B"  
/clone\_id="NIH\_MGC\_115"  
/note="Organ: pooled brain, lung, testis; Vector:  
pCMV-SPORT6; Site 1: NotI; Site 2: EcoRV (destroyed); RNA  
source anonymous pool of 6 male brains, age range 23-27; 1  
male lung, age 27, and 1 male testis, age 69. Library is  
oligo-dT primed and directionally cloned (EcoRV site is  
destroyed upon cloning). Average insert size 1.8 kb,  
insert size range 1-3 kb. Library is normalized and  
enriched for full-length clones and was constructed by C.  
Gruber (Invitrogen). Research Genetics tracking code  
021. Note: this is a NIH\_MGC library."  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,466-51 Length: 1038  
Score: 620.00 Matches: 126  
Percent Similarity: 79.64% Conservative: 7  
Best Local Similarity: 75.45% Mismatches: 21  
Query Match: 26.44% Indels: 13  
Gaps: 2  
DB: 5  
US-09-596-774-7 (1-443) x BM921412 (1-1038)  
QY 278 H15FROTHRG1YTHS6G1NPROG1NARPROG1UASP6Y5ARGPROARG1YSERVAL 297  
DB 86 CATCTCGAGGACAGCTTGGCCGATTACAGA----- 115  
QY 298 LYGGLYTHRG1YLEU-ASPHELEUG1UASP6Y5ARGPROARG1YSERVAL 317  
DB 116 ---GGCAGAGAGCTTGGCCGCTG---GATCCCAACTGCTACCTGCTGAGAGGAAAT 169  
QY 317 ELEUPHE1LE1YRG1VAL1LE1LE1THRALALEUTYRLEUARGALALYSPHESERARGSE 337  
DB 170 CCTCTTCA1TCATAGTGTCATTTCTAC1TGCCTTGTCTCTGAGAGGAAGTTCAAGAGAG 229  
QY 337 TALEGLU1THRALA1AENLEUG1NASP6Y5ARGPROARG1YSERVAL 357  
DB 230 CGCAGACGCCCGCGTACAGCAGGAGGCGAGAAACAGCTTAAACGAGCTCAATCTAGG 289

QY 357 YARGARG1UG1U1YTHASPVAL1LEUG1ULYLYARGAL1AARGASP6Y5ARGPRO1UMETG1YGI 377  
DB 290 ACGAAGAGAGAGTACGATGTTTGGACAAAGACGTCGCCGAGACCTTGAGATGGGAGG 349  
QY 377 YLYGGL1NARGARGARGAENP6Y5ARGPRO1UG1UG1VAL1YTRAPNALALEUG1NLYBAPPLY 397  
DB 350 AAACCCGACAGAGAGAAAGAAACCTTCAGAGAGGCTGTACAAATGACGAGAAAGATTA 409  
QY 397 SMEALAG1UALA1YRSE1G1U1LEGL1YTHLYEG1YGLUARGARG1YLYSG1YHI 417  
DB 410 GATGCGGAGGCTTACAGTACGATTTGGATGAAAGCGACGCGGAGGCGGAGGAGCA 469  
QY 417 SASEGLYLEUTYRG1NG1YLEUSERTH1ALATHLYASP6Y5THYTRAPNALALEUHSME 437  
DB 470 CGATGGCCTTACAGAGGTCCTCACTTACAGCAGCAGGACAGCAGCAGCCTTACAT 529  
QY 437 TGL1TH1LEU1A1APROARG 443  
DB 530 GCAGGCGCTTGGCCCTCGC 548  
RESULT 14  
LOCUS AL554350 1113 bp mRNA linear EST 30-MAR-2004  
DEFINITION AL554350 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA  
clone CSOD1082Y123 5-PRIME, mRNA sequence.  
ACCESSION AL554350  
VERSION AL554350.3 GI:45859108  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
REFERENCE 1 (bases 1 to 1113)  
AUTHORS Li, W.B., Gruber, C., Jeejee, J. and Polayes, D.  
TITLE Full-length cDNA libraries and normalization  
JOURNAL Unpublished (2001)  
COMMENT On Feb 15, 2001 this sequence version replaced gi:31276162.  
Genoscope - Centre National de Sequencage  
2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE  
Email: seqref@genoscope.cns.fr, Web: www.genoscope.cns.fr  
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime  
end enriched, double-strand cDNA was digested with Not I and cloned  
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library  
was normalized. Library was constructed by Life Technologies, a  
division of Invitrogen. This sequence belongs to sequence cluster  
484.x  
For more information about this cluster, see  
http://www.genoscope.cns.fr/cdnats-CSOD1082AE12QP1&c=484.r.  
FEATURES  
Source  
1. 1113  
Location/Qualifiers  
/organism="Homo sapiens"  
/mol\_type="mRNA"  
/db\_xref="taxon:9606"  
/clone="CSOD1082Y123"  
/tissue\_type="PLACENTA COT 25-NORMALIZED"  
/clone\_id="Homo sapiens PLACENTA COT 25-NORMALIZED"  
/note="1st strand cDNA was primed with a NotI-oligo(dT)  
primer. Five prime end enriched, double-strand cDNA was  
digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."  
ORIGIN  
Alignment Scores:  
Pred. No.: 3,82e-51 Length: 1113  
Score: 620.00 Matches: 126  
Percent Similarity: 79.64% Conservative: 7  
Best Local Similarity: 75.45% Mismatches: 21  
Query Match: 26.44% Indels: 13  
Gaps: 2  
DB: 1  
US-09-596-774-7 (1-443) x AL554350 (1-1113)

```

Oy 278 HisProthrglyThrsSerGlnProGlnArgProGluAspCysArgProArgGlySerVal 297
    |||||
Db 87 CATCTGCGAGGACAGATTGCCGATTACAGA----- 116
Oy 298 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyI 317
    |||||
Db 117 ---GGCACAGAGCTTTGGCTGCTG---GATCCAAACTGCTCTCCTGCTGATGATGAT 170
Oy 317 eleuPheIleTyrGlyValIleIleThraAlaLeuTyrLeuArgAlaLysPheSerArgSe 337
    |||||
Db 171 CCTCTCATCTATGATGATCTCTCATCTGCTGCTGCTGCTGAGTGAAGTTCAGCAGAG 230
Oy 337 rAlaGluThraAlaAlaAsnLeuGluAspProAsnGlnLeuTyrAsnGluLeuAsnLeuG 357
    |||||
Db 231 CGCAAGACGCCCCCGCTACCGACAGAGGCGCAGAACCACTATACGAGCTCAATCTAGG 290
Oy 357 YArgArgGluGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyI 377
    |||||
Db 291 ACGAAGAGAGAGTACGATGTTTGACAGAGAGCTGCGCGGACCTTGAGATGCGGGG 350
Oy 377 YLysGlnGlnArgArgAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspGly 397
    |||||
Db 351 AAAGCCGAGAGAGAGAGAACCTCAGAGAGGCTGTACATGAACTGCAGAAAGATGA 410
Oy 397 sMetAlaGluAlaTyrSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyYH 417
    |||||
Db 411 GATGCGGAGGCGCTACAGTGAATGGATGGAAGCGACGCGGAGGCGCAAGGGGCA 470
Oy 417 sAspGlyLeuTyrGlnGlyLeuSerThraAlaThrLysAspThrTyrAspAlaLeuHsMe 437
    |||||
Db 471 CGATGCGCTTACAGAGCTCTCAGTACAGCCACAGACACCTTACAGCGCTTCACAT 530
Oy 437 tGlnThrLeuAlaProArg 443
    |||||
Db 531 GCAGGCGCTGCCCTCGC 549

RESULT 15
LOCUS CR625818 1546 bp mRNA linear HTC 21-JUL-2004
DEFINITION full-length cDNA clone CS0DJ005YJ16 of T cells (Jurkat cell line)
ACCESSION CR625818
VERSION 1 GI:50506625
KEYWORDS HTC; CNSLT_cDNA.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 1546)
L1,W.B., Gruber,C., Jessee,J. and Polayes,D.
Full-length cDNA libraries and normalization
Unpublished
Contact: Feng Liang Email: fliang@lifetech.com URL:
http://fulllength.invitrogen.com/ Invitrogen Corporation 1600
Paraday Avenue
2 (bases 1 to 1546)
REFERENCE
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (20-JUL-2004) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail: seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
COMMENT
1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
end enriched, double-strand cDNA was digested with Not I and cloned
into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
was normalized. Library was constructed by Life Technologies, a
division of Invitrogen.
FEATURES
source location/Qualifiers
1..1546
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="CS0DJ005YJ16"
/tissue_type="T cells (Jurkat cell line) Cot

```

```

ORIGIN 10-normalized"
/plasmid="pCMVSPORT_6"
Alignment Scores:
Pred. No.: 6.08e-51 Length: 1546
Score: 620.00 Matches: 126
Percent Similarity: 79.64% Conservative: 7
Best Local Similarity: 75.45% Mismatches: 21
Query Match: 26.44% Indels: 13
DB: 3 Gaps: 2
US-09-596-774-7 (1-443) x CR625818 (1-1546)
Oy 278 HisProthrglyThrsSerGlnProGlnArgProGluAspCysArgProArgGlySerVal 297
    |||||
Db 69 CATCTGCGAGGACAGATTGCCGATTACAGA----- 98
Oy 298 LysGlyThrGlyLeu-AspPheLeuGluAspProLysLeuCysTyrLeuLeuAspGlyI 317
    |||||
Db 99 ---GGCACAGAGCTTTGGCTGCTG---GATCCAAACTGCTCTCCTGCTGATGATGAT 152
Oy 317 eleuPheIleTyrGlyValIleIleThraAlaLeuTyrLeuArgAlaLysPheSerArgSe 337
    |||||
Db 153 CCTCTCATCTATGATGATCTCTCATCTGCTGCTGCTGCTGAGTGAAGTTCAGCAGAG 212
Oy 337 rAlaGluThraAlaAlaAsnLeuGluAspProAsnGlnLeuTyrAsnGluLeuAsnLeuG 357
    |||||
Db 213 CGCAAGACGCCCCCGCTACCGACAGAGGCGCAGAACCACTATACGAGCTCAATCTAGG 272
Oy 357 YArgArgGluGluTyrAspValLeuGluLysLysArgAlaArgAspProGluMetGlyI 377
    |||||
Db 273 ACGAAGAGAGAGTACGATGTTTGACAGAGAGCTGCGCGGACCTTGAGATGCGGGG 332
Oy 377 YLysGlnGlnArgArgAsnProGlnGluGlyValTyrAsnAlaLeuGlnLysAspGly 397
    |||||
Db 333 AAAGCCGAGAGAGAGAGAACCTCAGAGAGGCTGTACATGAACTGCAGAAAGATGA 392
Oy 397 sMetAlaGluAlaTyrSerGluIleGlyThrLysGlyGluArgArgGlyLysGlyYH 417
    |||||
Db 393 GATGCGGAGGCGCTTACAGTGAATGGATGGAAGCGAGCGCGGAGGCGCAAGGGGCA 452
Oy 417 sAspGlyLeuTyrGlnGlyLeuSerThraAlaThrLysAspThrTyrAspAlaLeuHsMe 437
    |||||
Db 453 CGATGCGCTTACAGAGCTCTCAGTACAGCCACAGAGACACCTTACAGCGCTTCACAT 512
Oy 437 tGlnThrLeuAlaProArg 443
    |||||
Db 513 GCAGGCGCTGCCCTCGC 531

```

Search completed: April 20, 2005, 16:33:47  
Job time : 4354.69 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2005 Compugen Ltd.

OW protein - protein search, using bw model

Run on: April 20, 2005, 06:50:16 ; Search time 63.2157 Seconds  
(without alignments)  
3588.524 Million cell updates/sec

Title: US-09-596-774-7  
Perfect score: 2345  
Sequence: 1 QIOLVOSGPELKKRGEYTKI.....LSTATKOTYDALHMQTLAPR 443

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0  
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : UniProt\_03:\*  
1: uniprot\_sprot:\*  
2: uniprot\_trembl:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	904	38.6	241	2 Q921A6	Q921A6 mus musculu
2	832	35.5	244	2 Q65ZC8	Q65ZC8 mus musculu
3	778	33.2	240	2 Q65ZC9	Q65ZC9 mus musculu
4	776	33.1	255	2 Q6KB05	Q6KB05 mus musculu
5	767	32.7	243	2 Q7TQM2	Q7TQM2 mus musculu
6	746.5	31.8	218	2 Q925S1	Q925S1 mus musculu
7	744.5	31.7	298	2 Q9QYF0	Q9QYF0 synthetic c
8	733	31.3	487	2 Q65ZL2	Q65ZL2 mus sp. fv/
9	725.5	30.9	248	2 Q65Z07	Q65Z07 mus sp. b3/
10	713	30.4	164	1 P24161	P24161 mus musculu
11	618	26.4	164	1 P20963	P20963 mus musculu
12	605	25.8	206	1 P23020	P23020 mus musculu
13	602.5	25.7	163	1 Q9X619	Q9X619 sus scrofa
14	589.5	25.1	165	1 CD3Z_RABIT	CD3Z_RABIT
15	539	23.0	166	1 CD3Z_SHEEP	CD3Z_SHEEP
16	502.5	21.4	484	2 Q991A6	Q991A6 mus musculu
17	486	20.7	149	1 P01633	P01633 mus musculu
18	454.5	19.4	116	2 Q683Y7	Q683Y7 mus musculu
19	446.5	19.0	140	2 Q65ZL3	Q65ZL3 mus sp. c91
20	429.5	18.3	116	2 Q683Y8	Q683Y8 mus musculu
21	419	17.9	102	2 Q91L79	Q91L79 mus musculu
22	417	17.8	108	2 Q8V1U0	Q8V1U0 mus musculu
23	413	17.6	99	2 Q90L74	Q90L74 mus musculu
24	409	17.4	170	2 Q925S2	Q925S2 mus musculu
25	398	17.0	125	2 Q6P1L0	Q6P1L0 mus musculu
26	394.5	16.8	518	2 Q6N030	Q6N030 mus musculu
27	393	16.8	497	2 Q8VY24	Q8VY24 mus musculu
28	391.5	16.7	236	2 Q6GMX8	Q6GMX8 mus sapien
29	390	16.6	136	1 P01634	P01634 mus musculu
30	389	16.6	108	1 KVIY_HUMAN	KVIY_HUMAN
31	387.5	16.5	236	2 Q6P1H7	Q6P1H7 mus musculu

32	387.5	16.5	322	2 Q6KAV0	Q6KAV0 mus musculu
33	386.5	16.5	236	2 Q6GMW1	Q6GMW1 mus musculu
34	385.5	16.4	498	2 Q6N041	Q6N041 mus musculu
35	384	16.4	108	1 KVIY_HUMAN	KVIY_HUMAN
36	381.5	16.3	236	2 Q723Y4	Q723Y4 mus musculu
37	380	16.2	480	2 Q6P089	Q6P089 mus musculu
38	379	16.2	108	1 KVIY_HUMAN	KVIY_HUMAN
39	378	16.1	147	2 Q925S3	Q925S3 mus musculu
40	377.5	16.1	236	2 Q6GMX0	Q6GMX0 mus musculu
41	377	16.1	108	1 KVIY_HUMAN	KVIY_HUMAN
42	377	16.1	108	1 Q91L77	Q91L77 mus musculu
43	377	16.1	108	2 KVIY_HUMAN	KVIY_HUMAN
44	377	16.1	129	1 Q91L70	Q91L70 mus musculu
45	375	16.0	108	2 Q91L70	Q91L70 mus musculu

## ALIGNMENTS

### RESULT 1

ID	Q921A6	PRELIMINARY:	PRT:	241 AA.
AC	Q921A6	01-DEC-2001 (Tremblrel. 19, Created)		
DT	01-DEC-2001 (Tremblrel. 19, Last sequence update)			
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)			
DE	Anti-CEA 79 single chain Fv (Fragment).			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	MEDLINE=98170165; PubMed=9509426;			
RA	Chung J.H., Choi S.J., Kim H.J., Kim I.J., Choi I.H., Lee S.D.,			
RA	Yi K.S., Suh P.G., Ryu S.H., Chung H.K.;			
RT	"Cloning and characterization of cDNAs encoding VH and VL of a			
RT	monoclonal anti-CEA antibody (CEA 79) cross-reactive with NCA-95 and			
RT	generation of a single-chain Fv molecule (scfv)."			
RL	Mol. Cells 7:816-819 (1997).			
DR	EMBL; U88067; AAB48044.1; -.			
DR	PIR; S19965; S19965.			
DR	PIR; S19967; S19967.			
DR	PIR; S19968; S19968.			
DR	PIR; S26325; S26325.			
DR	HSSP; P01607; 18MW.			
DR	SMART; SM00406; IGV; 2.			
DR	PROSITE; PS50835; IG LIKE; 2.			
FT	NON_TER 1			
FT	NON_TER 241			
FT	SEQUENCE 241 AA; 26086 MW; 0276887248B9C771 CRC64;			

Query Match 38.6%; Score 904; DB 2; Length 241;  
Best Local Similarity 70.2%; Pred. No. 2.1e-58;  
Matches 172; Conservative 29; Mismatches 34; Indels 10; Gaps 4;

QY	1 QIOLVOSGPELKKRGEYTKISCKASGYPTNYGNWVQAQAGQGLKMWGINTSGESTF 60	QY	1 QIOLVOSGPELKKRGEYTKISCKASGYPTNYGNWVQAQAGQGLKMWGINTSGESTF 60
DB	1 QVRLQOSGPELKKRGEYTKISCKASGYPTDYGNWVQAQAGQGLKMWGINTTGEFTY 60	DB	1 QVRLQOSGPELKKRGEYTKISCKASGYPTDYGNWVQAQAGQGLKMWGINTTGEFTY 60
QY	61 ADDEKGRFDFSLSTANTAYLQINNLSKSDMATYFCARMEVYHGVYPMGCGTTVTVSSG 120	QY	61 ADDEKGRFDFSLSTANTAYLQINNLSKSDMATYFCARMEVYHGVYPMGCGTTVTVSSG 120
DB	61 ADDKGRFDFSLSTANTAYLQINNLSKSDMATYFCARKDLR-YFDWGGGTTVTVSSG 119	DB	61 ADDKGRFDFSLSTANTAYLQINNLSKSDMATYFCARKDLR-YFDWGGGTTVTVSSG 119
QY	121 GGSGGGGGGGGGSDIQLTQSHKFLSTVSDRVSITCKASQDVYNAVAWYQKQGQSK- 179	QY	121 GGSGGGGGGGGGSDIQLTQSHKFLSTVSDRVSITCKASQDVYNAVAWYQKQGQSK- 179
DB	120 GGSGGGGGGGGGSDIETQSSLSASLGKGYITCKASQDINKYAWQHKKGKPRS 179	DB	120 GGSGGGGGGGGGSDIETQSSLSASLGKGYITCKASQDINKYAWQHKKGKPRS 179
QY	180 ---LLIYSASRYGVSPFTGSGSGPDPTFTISSVOAEDLAVYFCQOHPTPTFGSGT 236	QY	180 ---LLIYSASRYGVSPFTGSGSGPDPTFTISSVOAEDLAVYFCQOHPTPTFGSGT 236
DB	180 AHTLHIT---QGITRSRSGSGSGRDYFSLISNLEBEDIATYYC-LHYDNLHTFGGCT 234	DB	180 AHTLHIT---QGITRSRSGSGSGRDYFSLISNLEBEDIATYYC-LHYDNLHTFGGCT 234
QY	237 KLEIK 241	QY	237 KLEIK 241

```

Db          235 KLEBK 239

RESULT 2
O65ZC8      PRELIMINARY; PRT; 244 AA.
AC 065ZC8;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Single-chain Fv (Fragment).
GN Name=scFv;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=97362799; PubMed=9219263;
RA Kontermann R.E., Wing M.G., Winter G.;
RT "Complement recruitment using bispecific diabodies.";
RL Nat. Biotechnol. 15:629-631(1997).
DR EMBL; Y13057; CAA73500.1; -.
DR InterPro; IPR003599; Ig_1.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; IG; 2.
DR PROSITE; PSS0835; IG_LIKE; 2.
FT NON_TER 1
FT NON_TER 244
SQ SEQUENCE 244 AA; 26127 MW; 4B1F17868338F2BF CRC64;

Query Match 35.5%; Score 832; DB 2; Length 244;
Best Local Similarity 61.9%; Pred. No. 4,1e-53;
Matches 151; Conservative 44; Mismatches 45; Indels 4; Gaps 2;

Oy 1 QIOLVOSGPELKKPGETVKISCKASGYPTNYGMNVKQAPGGLKMMGMINSTGSETF 60
Db 1 QVQLVQSGAEVKKPQDSVKVSCKASGYTFSDHYMHWVQAPGGLKMMGMINSTGSETF 60
Oy 61 ADPRKGRFDSLETSANTAYLQIINNLSKEDMATYFCAR--WEVYHGVPYWGQTTVTS 117
Db 61 AQRQGRVTMTDTSISAAYMEVSRRLRSDDTAVYYCARREGTSAIYG-MDVWGQGLVTV 119
Oy 118 SSGGSGSGGSGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVYNAVMYQKPGSG 177
Db 120 SSGGSGSGGSGGSDIQLTQSPSTLSASIGRVITTCASGSIYHMLAWYQKPGSKA 179
Oy 178 PKLIYSASRYTVGPSRFTSGSGSDPFTFTISSVQADLAVYFCQGHFRPTFTFGSGTK 237
Db 180 PKFLIYKASLASAPSRFSSGSGDTFTLTISLQPDPAFTYCCQGYSNVPLTFGGGTR 239
Oy 238 LEIK 241
Db 240 LEIK 243

RESULT 3
O65ZC9      PRELIMINARY; PRT; 240 AA.
AC 065ZC9;
DT 25-OCT-2004 (TREMBlrel. 28, Created)
DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
DE Single-chain Fv (Fragment).
GN Name=scFv;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
OX NCBI_TaxId=9606;

```

```

RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C14/77;
RX MEDLINE=97362799; PubMed=9219263;
RA Kontermann R.E., Wing M.G., Winter G.;
RT "Complement recruitment using bispecific diabodies.";
RL Nat. Biotechnol. 15:629-631(1997).
DR EMBL; Y13056; CAA73499.1; -.
DR InterPro; IPR003599; Ig_1.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003596; Ig_v.
DR Pfam; PF00047; Ig; 2.
DR SMART; SM00409; IG; 2.
DR PROSITE; PSS0835; IG_LIKE; 2.
FT NON_TER 1
FT NON_TER 240
SQ SEQUENCE 240 AA; 25569 MW; FDCFD3454F64B373 CRC64;

Query Match 33.2%; Score 778; DB 2; Length 240;
Best Local Similarity 60.3%; Pred. No. 3.7e-49;
Matches 146; Conservative 39; Mismatches 53; Indels 4; Gaps 2;

Oy 1 QIOLVOSGPELKKPGETVKISCKASGYPTNYGMNVKQAPGGLKMMGMINSTGSETF 60
Db 1 QVQLVQSGGLVQPGSIRLSCAASGFTSSGYMHWVQAPGGLKMMGMINSTGSETF 60
Oy 61 ADPRKGRFDSLETSANTAYLQIINNLSKEDMATYFCAR--WEVYHGVPYWGQTTVTS 119
Db 61 ADSVKGKFTISRNSKNTLLQWNSLRABDTAVYYCARDM--GDSLDPWGKQTLTVTS 117
Oy 120 GGGSGSGGSGGSDIQLTQSHKFLSTSVGRVSIITCKASQDVYNAVMYQKPGSPK 179
Db 118 GGGSGSGGSGGSDIQLTQSPSTLSASIGRVITTCASGSIYHMLAWYQKPGSKA 177
Oy 180 LLIYSASRYTVGPSRFTSGSGSDPFTFTISSVQADLAVYFCQGHFRPTFTFGSGTKLE 239
Db 178 LLIYKASLASAPSRFSSGSGDTFTLTISLQPDPAFTYCCQGYSNVPLTFGGGTR 237
Oy 240 IK 241
Db 238 IK 239

RESULT 4
O6KB05      PRELIMINARY; PRT; 255 AA.
AC 06KB05;
DT 05-JUL-2004 (TREMBlrel. 27, Created)
DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
DE scFv B8E5 protein (Fragment).
GN Name=scFv B8E5;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxId=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Bald/c;
RA Peter J.C., Wallukat G., Tugler J., Maurice D., Reegel J.C.,
RA Briand J.P., Hoebeke J.;
RL Submitted (JUN-2004) to the EMBL/Genbank/DBJ databases.
DR EMBL; AJ746180; CAG34081.1; -.
DR HSSP; P01837; 1KCR.
DR InterPro; IPR003599; Ig_1.
DR InterPro; IPR007110; Ig_1like.
DR InterPro; IPR003596; Ig_v.
DR SMART; SM00409; IG; 2.
DR PROSITE; PSS0835; IG_LIKE; 2.
FT NON_TER 1
FT NON_TER 255
SQ SEQUENCE 255 AA; 27445 MW; B68BD38395DF713B CRC64;

```



Query Match 33.1%; Score 776; DB 2; Length 255;  
 Best Local Similarity 59.8%; Pred. No. 5, 7e-49;  
 Matches 149; Conservative 36; Mismatches 56; Indels 8; Gaps 2;

QY 1 QIOLVSGPELKKPGETYKISCKASGYPTNYGMNWKQAPGQGLKMMGMINSTGSESTF 60  
 DB 1 QVOLIQQSGSDIYKPCGSLKVSCKASGFTFSSYGMWVQTPDKRLBVALTISGGSYTY 60  
 QY 61 ADDEKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEYHGVYVWGQGITVTVSSG 118  
 DB 61 PDSVAGRTTISDNKNTLYLQMSLKSBDTAMMYCAHINRYDGAFDYWGQGITVTVSS 120  
 QY 119 SGGSGSGSGSGSGSDIQLTQSHKFLSTSVGDRAVITCKASQDVYNA-----VAATYQQ 172  
 DB 121 SGGSGSGSGSGSGSDIYVMAQPSLSVSADEKYMCKSSQSLNBRNKNTYLAAYQQ 180  
 QY 173 KPGQSPKLLIYASASRYTGVPSRFTGSGSDPFTFTISSVQAEDLAVYPCQHFPTPTF 232  
 DB 181 KPGQSPKLLIYASASRYTGVPSRFTGSGSDPFTFTISSVQAEDLAVYPCQNDHSYPLTF 240  
 QY 233 GSGTKLEIK 241  
 DB 241 GAGTKLEIK 249

RESULT 5

Q7QOM2 PRELIMINARY; PRT; 243 AA.  
 AC Q7QOM2;  
 DT 01-OCT-2003 (TREMBLrel. 25, Created)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last sequence update)  
 DT 01-MAR-2004 (TREMBLrel. 26, Last annotation update)  
 DE Scfv 6H8 protein (Fragment).  
 GN Name=Scfv 6H8;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RX MEDLINE=22853226; PubMed=12860977; DOI=10.1074/jbc.M306877200;  
 RA Peter J.C., Eftekhari P., Billiard P., Wallukke G., Hoebeke J.;  
 RT "scfv single chain antibody variable fragment as inverse agonist for  
 the beta-2 adrenergic receptor."  
 RL J. Biol. Chem. 278:36740-36747(2003).  
 DR EMBL; AJ574851; CAB00495.1; -.  
 DR HSBP; P01751; IAGW.  
 DR InterPro; IPR007110; IG-like.  
 DR InterPro; IPR003596; IG\_v.  
 DR SMART; SM00406; IGV; 2.  
 DR PROSITE; PS50835; IG-LIKE; 2.  
 FT NON TER  
 FT PROSITE 1  
 SQ SEQUENCE 243 AA; 25976 MW; BEPFF642DCFAF76 CRC64;

Query Match 32.7%; Score 767; DB 2; Length 243;  
 Best Local Similarity 60.6%; Pred. No. 2, 4e-48;  
 Matches 146; Conservative 42; Mismatches 49; Indels 4; Gaps 3;

QY 1 QIOLVSGPELKKPGETYKISCKASGYPTNYGMNWKQAPGQGLKMMGMINSTGSESTF 60  
 DB 1 QVOLIQQSGSELVPRGASVYKLSCKASGYFTTYMMWAKORHGQGLKMMGMINSTGSESTF 60  
 QY 61 ADDEKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEYHGVYVWGQGITVTVSSG 120  
 DB 61 DEKFNKILTVDTSSSTAYVHLSLASEDSAVYYCARG--RG-LDVWAGAGTTLTVSSG 117  
 QY 121 GGGSGSGSGSGSDIQLTQSHKFLSTSVGDRAVITCKASQDVYNAAYVQKQSPKL 180  
 DB 118 GGGSGSGSGSGSDIQLTQSHKFLSTSVGDRAVITCKASQDVYNAAYVQKQSPKL 177  
 QY 181 LIYASASRYTGVPSRFTGSGSDPFTFTISSVQAEDLAVYPCQHFPTPTFSGTKLEI 240

DB 178 LISGATSIETGVPSRFTGSGSGKDYTLSTLQTEDVATYVCOQWSTR-TFGGGTKLEI 236  
 QY 241 K 241  
 DB 237 K 237

RESULT 6

Q925S1 PRELIMINARY; PRT; 218 AA.  
 AC Q925S1;  
 DT 01-DEC-2001 (TREMBLrel. 19, Created)  
 DT 01-DEC-2001 (TREMBLrel. 19, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE MRE5 (Fragment).  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 NX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RX PubMed=11819679;  
 RA Cui D., Zeng G., Yan X., Wang F., Tian F., Ren D., Zhao T., Li X.,  
 Su C.;  
 RT "Mechanism of exogenous nucleic acids and their precursors improving  
 the repair of intestinal epithelium after irradiation in mice."  
 RL World J. Gastroenterol. 6:709-717(2000).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=BALB/c;  
 RA Cui D., Zeng G., Yan X., Li X., Su C.;  
 RT "Cloning of mouse genes related to repairing of intestinal epithelium  
 of the irradiated mice by treatment with the intestinal RNA of mice of  
 the same strain."  
 RL Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med. 19:71-80(2001).  
 DR EMBL; AF240168; AAK43733.1; -.  
 DR HSBP; P01665; IONZ.  
 DR SMART; SM00406; IGV; 1.  
 DR PROSITE; PS50835; IG-LIKE; 1.  
 FT NON TER  
 FT PROSITE 218  
 SQ SEQUENCE 218 AA; 23013 MW; 527E4FA8F7982817 CRC64;

Query Match 31.8%; Score 746.5; DB 2; Length 218;  
 Best Local Similarity 66.2%; Pred. No. 6, 8e-47;  
 Matches 141; Conservative 28; Mismatches 39; Indels 5; Gaps 2;

QY 1 QIOLVSGPELKKPGETYKISCKASGYPTNYGMNWKQAPGQGLKMMGMINSTGSESTF 60  
 DB 3 QVOLIQQSGPELKKPGETYKISCKASGYFTTYMMWAKORHGQGLKMMGMINSTGSESTF 62  
 QY 61 ADDEKGRFDSLETSANTAYLQINNLKSEDMATYFCARMEYHGVYVWGQGITVTVSSG 120  
 DB 63 ABEFKGRFAFSLETSANTAYLQINNLKSEDMATYFCARMEYHGVYVWGQGITVTVSSG 121  
 QY 121 GGGSGSGSGSGSDIQLTQSHKFLSTSVGDRAVITCKASQDVYNA-----VAATYQQPQ 176  
 DB 122 GGGSGSGSGSGSDIQLTQSHKFLSTSVGDRAVITCKASQDVYNAAYVQKQSPKL 181  
 QY 177 SPKLLIYASASRYTGVPSRFTGSGSDPFTFTISSVQAEDLAVYPCQHFPTPTFSGTKLEI 240  
 DB 182 PKLLIYASASRYTGVPSRFTGSGSDPFTFTISSVQAEDLAVYPCQHFPTPTFSGTKLEI 240

RESULT 7

Q9QYF0 PRELIMINARY; PRT; 298 AA.  
 AC Q9QYF0;  
 DT 01-MAY-2000 (TREMBLrel. 13, Created)  
 DT 01-MAY-2000 (TREMBLrel. 13, Last sequence update)  
 DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)  
 DE CN 8 single chain antibody.



Db 3 VKLVESGGGLVPGGSLKSLSCATSGFTESDYVWYWRQTPKRLKLEWAVYISNDSSAAYS 62  
 Qy 62 DDFGRPFSLSTANTAYLQINNKSRDMATPCARFWEVHGVPYWGQGTTLTVSSGG 121  
 Db 63 DTVGKRFITSDNANNTYLQWSRKSEDTALYSARGLWAGMAYWGQGLTVVSSGG 122  
 Qy 122 GSGGGGGGGGGGDIQTLQSHKFLSTSVQDRVSTTCASODPVYNA-----VAMVQOKPQ 176  
 Db 123 GSGGGGGGGGGGSDVLTMTQSPSLFVSLGDDQASISCRSQIIIVHSNGMTYLEWYLQKPGQ 182  
 Qy 177 SPKLLIYSASSRYGVPSRFRTGSGGPPFTTISVQADLAVYFCQGHFTPTFGSGT 236  
 Db 183 SPKLLIYKVSNNRFGVPPRFSGSGSGGTFTLKSIRVEADLGVYVCFQGSHPFTFGSGT 242  
 Qy 237 KLEIKR 242  
 Db 243 KLEIKR 248

RESULT 10  
 CD3Z\_MOUSE STANDARD; PRT; 164 AA.  
 ID CD3Z\_MOUSE  
 AC P24161; Q9D3G3;  
 DT 01-MAR-1992 (Rel. 21, Created)  
 DT 01-MAR-1992 (Rel. 21, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor  
 T3 zeta chain).  
 GN Name=CD3z; Synonyma=Tcrz;  
 OS Mus musculus (Mouse).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 CC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
 OX NCBI\_TaxID=10090;  
 RN [1]  
 RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.  
 RX MEDLINE=86145643; PubMed=3278377;  
 RA Weissman A.M., Banayash M., Hou D., Samelson L.E., Burgess W.H.,  
 RA Klausner R.D.;  
 RT "Molecular cloning of the zeta chain of the T cell antigen receptor.";  
 RL Science 239:1018-1021 (1988).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC TISSUE=Liver;  
 RX MEDLINE=89327299; PubMed=2787796;  
 RA Banayash M., Hsu V.W., Seldin M.F., Klausner R.D.;  
 RT "The isolation and characterization of the murine T cell antigen  
 receptor zeta chain gene.";  
 RL J. Biol. Chem. 264:13252-13257 (1989).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6J; TISSUE=Thymus;  
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;  
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi Y., Bono H., Kondo S.,  
 RA Nakajima I., Osato N., Saito R., Suzuki H., Yamataka I., Kiyosawa H.,  
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,  
 RA Baidetelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,  
 RA Schirral L.M., Kanapin A., Matwuda H., Batalov S., Beisel K.W.,  
 RA Blake J.A., Brad D., Brusic V., Chochia C., Corbani L.E., Cousins S.,  
 RA Dalla E., Dragan T.A., Fletcher C.F., Forrest A., Frazer K.S.,  
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,  
 RA Grimmond S., Gusninch S., Hirokawa N., Jackson I.J., Jarvis E.D.,  
 RA Kanai A., Kawai H., Kawasawa Y., Kedzierski R.M., King B.L.,  
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,  
 RA Maglott D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,  
 RA Nagasima T., Numata K., Okido T., Pavan W.J., Petosa G., Pesole G.,  
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,  
 RA Raveai T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Rinaldi M.,  
 RA Savelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,  
 RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita M.,  
 RA Verardo R., Wagner L., Wahlested C., Wang Y., Watanabe Y., Wells C.,  
 RA Wilming L.G., Wyshew-Borls A., Yanagisawa M., Yang I., Yang L.,  
 RA Yuan Z., Zavolan M., Zhu Y., Zimmer A., Carlini P., Hayatsu N.,  
 RA Hirozane-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,  
 RA Hara A., Hashizume W., Imocani K., Ishii Y., Itoh M., Kagawa I.,  
 RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shinagawa A.,  
 RA Yasunishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,  
 RA Birney E., Hayashizaki Y.;  
 RT "Analysis of the mouse transcriptome based on functional annotation of  
 60,770 full-length cDNAs.";  
 RL Nature 420:563-573 (2002).  
 RN [4]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=C57BL/6; TISSUE=Hematopoietic;  
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;  
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,  
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,  
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,  
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,  
 RA Brownstein M.J., Ueclin T.B., Toshlyuk S., Carninci P., Prange C.,  
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,  
 RA Bosak S.A., McGowan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,  
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulik S.W.,  
 RA Villalón D.K., Wuzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
 RA Fahy J., Hellon E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,  
 RA Whitting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,  
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,  
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,  
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;  
 RT "Generation and initial analysis of more than 15,000 full-length human  
 and mouse cDNA sequences.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).  
 RN [5]  
 RP PARTIAL SEQUENCE.  
 RX MEDLINE=90239005; PubMed=2139725;  
 RA Jin J.Y., Clayton L.K., Howard F.D., Koyasu S., Sieh M.,  
 RA Steinbrich R., Tarr G.B., Reinherz E.L.;  
 RT "Molecular cloning of the CD3 zeta subunit identifies a CD3 zeta-  
 related product in thymus-derived cells.";  
 RL Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323 (1990).  
 RN [6]  
 RP ALTERNATIVE SPLICING.  
 RX PubMed=2150596;  
 RA Ohno H., Saito T.;  
 RT "CD3 zeta and eta chains are produced by alternative splicing from a  
 common gene.";  
 RL Int. Immunol. 2:1117-1119 (1990).  
 RN [7]  
 RP ERRATUM.  
 RA Ohno H., Saito T.;  
 RL Int. Immunol. 4:1339-1339 (1992).  
 RN [8]  
 RP INTERACTION WITH SLA.  
 RX MEDLINE=20130290; PubMed=10662792;  
 RA Sosinowski T., Pandey A., Dixit V.M., Weiss A.;  
 RT "Src-like adaptor protein (SLAP) is a negative regulator of T cell  
 receptor signaling.";  
 RL J. Exp. Med. 191:463-474 (2000).  
 RN [9]  
 RP INTERACTION WITH SLA2.  
 RX MEDLINE=2201997; PubMed=11891219; DOI=10.1074/jbc.M110318200;  
 RA Pandey A., Ibarrola N., Kratchmarova I., Fernandez M.M.,  
 RA Constantinescu S.N., Ohara O., Sawadikoso S., Lodish H.F., Mann M.,  
 RA "A novel Src homology 2 domain-containing molecule, Src-like adapter  
 protein-2 (SLAP-2), which negatively regulates T cell receptor  
 signaling.";  
 RL J. Biol. Chem. 277:19131-19138 (2002).  
 CC -I- FUNCTION: Probable role in assembly and expression of the TCR  
 complex as well as signal transduction upon antigen triggering.  
 CC -I- SUBUNIT: Interacts with DOK2 (by similarity). The TCR/CD3 complex  
 of T lymphocytes consists of either a TCR alpha/beta or TCR  
 gamma/delta heterodimer coexpressed at the cell surface with the  
 invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and

ID	CD3ζ_HUMAN	STANDARD	PRT	164 AA
DT	01-FEB-1991 (Rel. 17, Created)			
DT	28-FEB-2003 (Rel. 41, Last sequence update)			
DT	25-OCT-2004 (Rel. 45, Last annotation update)			
DE	T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor T3 zeta chain).			
DR	EMBL; M19729; AAAA0171.1; -			
DR	EMBL; J04967; AAA50301.1; -			
DR	EMBL; AK017904; BAB30997.1; -			
DR	EMBL; BC052824; AAH52824.1; -			
DR	PIR; A40104; A40104.			
DR	MGI; MGI:88334; Cd3z.			
DR	InterPro; IPR003110; ITAM.			
DR	Pfam; PF02189; ITAM; 3.			
KM	Alternative splicing; Direct protein sequencing; Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.			
FT	SIGNAL	1	21	
FT	CHAIN	22	164	
FT	DOMAIN	22	30	
FT	TRANSMEM	31	51	
FT	DOMAIN	52	164	
FT	DOMAIN	69	89	
FT	DOMAIN	108	129	
FT	DOMAIN	139	159	
FT	DISULFID	32	32	
FT	MOD RES	153	153	
FT	CONFLICT	153	153	
SO	SEQUENCE	164 AA; 18637 MW; 186022035A312831 CRC64;		
QY	Query Match	30.4%; Score 713; DB 1; Length 164;		
Db	Best local Similarity	99.3%; Pred. No. 1.4e-44;		
Db	Matches 138; Conservative	0; Mismatches 1; Indels 0; Gaps 0;		
QY	305 LEDPKLCVLDGILFIYGVITITALLYLRKFSRSATFANLQDPNLYNELNGRREEDV	364		
Db	26 LLDPLCYLLDGILFIYGVITITALLYLRKFSRSATFANLQDPNLYNELNGRREEDV	85		
QY	365 LEKRRADPEWGGQQRNRNPQEGVYNALQDKMAEVAEISIGTKERRRGKGHDLVQGL	424		
Db	86 LEKRRADPEWGGQQRNRNPQEGVYNALQDKMAEVAEISIGTKERRRGKGHDLVQGL	145		
QY	425 STATKDTYDALHMQTLAPR	443		
Db	146 STATKDTYDALHMQTLAPR	164		
RESULT 11				
CD3ζ_HUMAN				
CD3ζ_HUMAN	STANDARD	PRT	164 AA	
P20963; Q8TAX4;				
01-FEB-1991 (Rel. 17, Created)				
28-FEB-2003 (Rel. 41, Last sequence update)				
25-OCT-2004 (Rel. 45, Last annotation update)				
T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor T3 zeta chain).				
Name=CD3ζ; Synonyms=T3ζ, TCRζ;				

OS Homo sapiens (Human).  
OC Eumetazoa: Metazoa: Chordata: Craniata: Vertebrata: Euteleostomi;  
OC Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.  
OC NCBI\_TaxID=9606;  
[1]  
RN  
RP SEQUENCE FROM N.A.  
RX MEDLINE=89071765; PubMed=2974162;  
RA Weissman A.M., Hou D., Orlloff D.G., Modi W.S., Seunanez H.,  
RT "Molecular cloning and chromosomal localization of the human T-cell  
RT receptor zeta chain: distinction from the molecular CD3 complex.";  
RL Proc. Natl. Acad. Sci. U.S.A. 85:9709-9713(1988).  
[2]  
RN  
RP SEQUENCE FROM N.A.  
RC TISSUE=Pancreas;  
RX MEDLINE=2238825; PubMed=12477932; DOI=10.1073/pnas.242603899;  
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,  
RA Klausner R.D., Collins F.S., Wagner L., Shemmen C.M., Schuler G.D.,  
RA Altschul S.F., Zeeberg B., Buacow K.H., Scheffer C.F., Bhat N.K.,  
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,  
RA Diachenko L., Marushina K., Farmer A.S., Rubin G.M., Hong L.,  
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,  
RA Brownstein M.J., Usdin T.B., Petrosky J.S., Carrinco P., Prange C.,  
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,  
RA Bosak S.S., McGwan P.J., McKernan K.J., Malek U.A., Gunaratne P.H.,  
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,  
RA Villalon D., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,  
RA Paley J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,  
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,  
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson W.C.,  
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,  
RA Butcherfield Y.S.N., Krzywinski M.I., Skalska U., Smalley D.E.,  
RA Schmechel A., Schein J.E., Jones S.J.M., Marra M.A.;  
RT "Generation and initial analysis of more than 15,000 full-length human  
RT and mouse cDNA sequences.";  
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).  
[3]  
RN  
RP INTERACTION WITH HIV-1 NEF.  
RX PubMed=10224289;  
RA Xu X.-N., Laffert B., Screaton G.R., Kraft M., Wolf D., Kolaas W.,  
RA Mongkolkeapay J., McMichael A.J., Baur A.S.;  
RT "Induction of Fas ligand expression by HIV involves the interaction of  
RT Nef with the T cell receptor zeta chain.";  
RL J. Exp. Med. 189:1489-1496(1999).  
[4]  
RN  
RP INTERACTION WITH SLA.  
RX MEDLINE=9398055; PubMed=10449770; DOI=10.1073/pnas.96.17.9775;  
RA Tang J., Sawaedikosol S., Chang J.-H., Burckoff S.J.;  
RT "SLAP, a dimeric adapter protein, plays a functional role in T cell  
RT receptor signaling.";  
RL Proc. Natl. Acad. Sci. U.S.A. 96:9775-9780(1999).  
[5]  
RN  
RP INTERACTION WITH CDCK2.  
RX MEDLINE=22165501; PubMed=12176041; DOI=10.1016/S0006-291X(02)00931-2;  
RA Nishihara H., Maeda M., Tsuda M., Makino Y., Sawa H., Nagashima K.,  
RA Tanaka S.;  
RT "CDCK2 mediates T cell receptor-induced activation of Rac2 and IL-2  
RT transcription.";  
RL Biochem. Biophys. Res. Commun. 236:716-720(2002).  
[6]  
RN  
RP PHOSPHORYLATION SITES TYR-83; TYR-111; TYR-123 AND TYR-142.  
RX PubMed=12522270; DOI=10.1073/pnas.2436191100;  
RA Salomon A.R., Ficarot S.B., Brill L.M., Brinker A., Phung Q.T.,  
RA Ericson C., Sauer K., Brock A., Horn D.M., Schultz P.G., Peters E.C.;  
RT "Profiling of tyrosine phosphorylation pathways in human cells using  
RT mass spectrometry.";  
RL Proc. Natl. Acad. Sci. U.S.A. 100:443-448(2003).  
[7]  
RN  
RP STRUCTURE BY NMR OF 136-149.  
RX MEDLINE=93010600; PubMed=7680960; DOI=10.1016/S0092-8674(93)90405-F;  
RA Makam G., Shoelson S.B., Pant N., Cowburn D., Kurtyan J.;  
RT "Binding of a high affinity phosphotyrosyl peptide to the Src SH2  
RT domain: crystal structures of the complexed and peptide-free forms.";  
RL Proc. Natl. Acad. Sci. U.S.A. 90:1015-1019(1993).

```

RL Cell 72:779-790(1993).
CC -1- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
CC cell surface with the invariant subunits of CD3 labeled gamma,
CC delta, epsilon, zeta, and eta. CD3-zeta forms either homodimers or
CC heterodimers with CD3-eta. Interacts with SIA and SIA2. Interacts
CC with DOCK2. Interacts with HIV-1 Nef protein.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=CD-3-zeta;
CC IsoId=P20963-1; Sequence=Displayed;
CC Name=CD-3-eta;
CC IsoId=P20963-2; Sequence=Not described;
CC -1- PTM: Phosphorylated on Tyr residues after T-cell receptor
CC triggering (By similarity).
CC -1- SIMILARITY: Belongs to the CD32/FCER1G family.
CC -1- SIMILARITY: Contains 3 ITAM domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL: J04132; AAA60394.1; -.
DR EMBL: BC025703; AAH25703.1; -.
DR PIR: A31768; A31768.
DR PDB: 1TCE; NMR: B-136-149.
DR Genev: HGNC:1677; CD3Z.
DR H-InvDB: HIX0001296; -.
DR MIM: 186780; -.
DR GO: GO:0005886; C:Plasma membrane; TAS.
DR GO: GO:0042101; C:T-cell receptor complex; TAS.
DR GO: GO:0005515; F:Protein binding; IPI.
DR GO: GO:0042803; F:Protein homodimerization activity; NAS.
DR InterPro: IPR003110; ITAM.
DR Pfam: PF02189; ITAM; 3.
DR SMART: SM00077; ITAM; 3.
DR 3D-structure: Alternative splicing; Phosphorylation; Receptor; Repeat;
KM Signal; T-cell; Transmembrane.
FT SIGNAL 1 21
FT CHAIN 22 164
FT DOMAIN 22 30 T-cell surface glycoprotein CD3 zeta
FT TRANSSEM 31 51 chain.
FT DOMAIN 52 164 Extracellular (Potential).
FT DOMAIN 69 89 Potential.
FT DOMAIN 108 129 Cytoplasmic (Potential).
FT DOMAIN 139 159 ITAM 1.
FT DISUPID 32 32 ITAM 2.
FT MOD_RES 83 83 ITAM 3.
FT MOD_RES 83 83 Interchain (Potential).
FT MOD_RES 111 111 Phosphotyrosine.
FT MOD_RES 123 123 Phosphotyrosine.
FT MOD_RES 142 142 Phosphotyrosine.
FT MOD_RES 153 153 Phosphotyrosine (By similarity).
FT CONFLICT 60 61 DA -> EP (in Ref. 1).
FT CONFLICT 101 101 Missing (in Ref. 1).
SQ SEQUENCE 164 AA; 18696 MW; 940826034856EE3 CMC64;
Query Match 26.4%; Score 618; DB 1; Length 164;
Best Local Similarity 84.9%; Pred. No. 1.3e-37;
Matches 118; Conservative 7; Mismatches 14; Indels 0; Gaps 0;

```

```

DB 86 LDKRGADPEWGKRPQRRNPOEGVNALQDKMAEAVSEIGTGERRRGKHGDLVQGL 145
OY 425 STATQTYDALHMQTLAPR 443
DB 146 STATQTYDALHMQALPPR 164

RESULT 12
CD3H_MOUSE STANDARD; PRT; 206 AA.
ID CD3H_MOUSE
AC P29020;
DT 01-DEC-1992 (Rel. 24, Created)
DT 01-DEC-1992 (Rel. 24, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE T-cell surface glycoprotein CD3 eta chain precursor (T-cell receptor
DE T3 eta chain).
GN Name=CD3z; Synonyms=Cd3h;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
CX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., AND PARTIAL SEQUENCE.
RX MEDLINE=90239005; PubMed=2139725;
RA Jin Y.J., Clayton L.K., Howard F.D., Koyasu S., Sieh M.,
RA Reinherz E.L., Tarr G.E., Reinherz E.L.;
RT "Molecular cloning of the CD3 eta subunit identifies a CD3 zeta-
RT related product in thymus-derived cells.";
RT Proc. Natl. Acad. Sci. U.S.A. 87:3319-3323(1990).
RN [2]
RP SEQUENCE OF 144-206 FROM N.A.
RX MEDLINE=91271358; PubMed=1828894;
RA Clayton L.K., D'Adamo L., Sieh M., Hussey R.E., Koyasu S.,
RA Reinherz E.L., Howard F.B.;
RT "CD3 eta and CD3 zeta are alternatively spliced products of a common
RT genetic locus and are transcriptionally and/or post-transcriptionally
RT regulated during T-cell development.";
RT Proc. Natl. Acad. Sci. U.S.A. 88:5202-5206(1991).
RN [3]
RP SEQUENCE OF 144-206 FROM N.A.
RX PubMed=2150596;
RA Ohno H., Saito T.;
RT "CD3 zeta and eta chains are produced by alternative splicing from a
RT common gene.";
RT Int. Immunol. 2:1117-1119(1990).
RN [4]
RP ERRATUM.
RA Ohno H., Saito T.;
RL Int. Immunol. 4:1339-1339(1992).
CC -1- FUNCTION: Probable role in assembly and expression of the TCR
CC complex as well as signal transduction upon antigen triggering.
CC -1- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a
CC TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the
CC cell surface with the invariant subunits of CD3 labeled gamma,
CC delta, epsilon, zeta, and eta. CD3-eta can be complexed in a
CC heterodimeric form with CD3-zeta subunit. CD3-eta homodimer has
CC not been observed.
CC -1- SUBCELLULAR LOCATION: Type I membrane protein.
CC -1- ALTERNATIVE PRODUCTS:
CC Event-Alternative splicing; Named isoforms=2;
CC Name=CD-3-eta;
CC IsoId=P29020-1; Sequence=Displayed;
CC Name=CD-3-zeta;
CC IsoId=P29020-1; Sequence=External;
CC -1- SIMILARITY: Belongs to the CD32/FCER1G family.
CC -1- SIMILARITY: Contains 3 ITAM domains.
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial

```

entitles requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; M33158; AAA37398.1; -  
 DR EMBL; M76711; AAA40403.1; -  
 DR PIR; A35900; A35900.  
 DR MGD; MGI:88334; Cd3z.  
 DR InterPro; IPR003110; ITAM.  
 DR Pfam; PF02189; ITAM; 2.  
 DR SMART; SM00077; ITAM; 2.  
 DR SMART; SM00077; ITAM; 2.  
 KM Alternative splicing; Direct protein sequencing; Receptor; Repeat;  
 FT Signal; T-cell; Transmembrane.  
 FT CHAIN 1 21  
 FT SIGNAL 1 21  
 FT CHAIN 22 206  
 FT SIGNAL 22 206  
 FT TRANSMEM 22 30  
 FT DOMAIN 31 51  
 FT DOMAIN 52 206  
 FT DOMAIN 69 89  
 FT DOMAIN 108 129  
 FT DOMAIN 139 159  
 FT DISULFID 32 32  
 FT DISULFID 32 32  
 SO SEQUENCE 206 AA; 23339 MW; 829256A2CF444444 CRC64;

Query Match 25.8%; Score 605; DB 1; Length 206;  
 Best Local Similarity 99.2%; Pred. No. 1.6e-36;  
 Matches 117; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 305 LEDPKCYLDDGLFTYGVITITLYLRKFSRAETANLDDPNQLYNELNGRREYDV 364  
 DB 26 LDDPKCYLDDGLFTYGVITITLYLRKFSRAETANLDDPNQLYNELNGRREYDV 85

QY 365 LEKRRARDPEMGKQOQRRRNPOEGVYNALQKDKMAEYSEIGTKGERRRGKHGDLGYQ 422  
 DB 86 LEKRRARDPEMGKQOQRRRNPOEGVYNALQKDKMAEYSEIGTKGERRRGKHGDLGYQ 143

RESULT 13  
 CD32\_PIG STANDARD; PRT; 163 AA.  
 AC O9XST9;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor T3 zeta chain).  
 GN Name=CD32;  
 OS Sus scrofa (Pig).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Cetartiodactyla; Suina; Suidae; Sus.  
 OX NCBI\_TaxID=9823;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Minnesota miniature swine;  
 RA Jie H.-B., Yim D., Kim Y.B.;  
 RT "The molecular cloning of porcine CD3 zeta.";  
 RL Submitted (May-1999) to the EMBL/GenBank/DBJ databases.  
 CC -I- FUNCTION: Probable role in assembly and expression of the TCR complex as well as signal transduction upon antigen triggering.  
 CC -I- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the cell surface with the invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and eta (By similarity).  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- PTM: Phosphorylated on Tyr residues after T-cell receptor triggering (By similarity).  
 CC -I- SIMILARITY: Belongs to the CD32/FCER1G family.  
 CC -I- SIMILARITY: Contains 3 ITAM domains.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way

modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

-----

DR EMBL; AF153830; AAD34640.1; -  
 DR InterPro; IPR003110; ITAM.  
 DR Pfam; PF02189; ITAM; 3.  
 DR SMART; SM00077; ITAM; 3.  
 KM Phosphorylation; Receptor; Repeat; Signal; T-cell; Transmembrane.  
 FT Signal 1 21  
 FT SIGNAL 1 21  
 FT CHAIN 22 163  
 FT SIGNAL 22 163  
 FT CHAIN 22 163  
 FT TRANSMEM 31 51  
 FT DOMAIN 52 163  
 FT DOMAIN 69 89  
 FT DOMAIN 107 128  
 FT DOMAIN 138 158  
 FT DISULFID 32 32  
 FT DISULFID 152 152  
 FT MOD RES 152 152  
 SO SEQUENCE 163 AA; 18568 MW; 34898620B67167C7 CRC64;

Query Match 25.7%; Score 602.5; DB 1; Length 163;  
 Best Local Similarity 84.2%; Pred. No. 1.8e-36;  
 Matches 117; Conservative 7; Mismatches 14; Indels 1; Gaps 1;

QY 305 LEDPKCYLDDGLFTYGVITITLYLRKFSRAETANLDDPNQLYNELNGRREYDV 364  
 DB 26 LDDPKCYLDDGLFTYGVITITLYLRKFSRAETANLDDPNQLYNELNGRREYDV 85

QY 365 LEKRRARDPEMGKQOQRRRNPOEGVYNALQKDKMAEYSEIGTKGERRRGKHGDLGYQ 424  
 DB 86 LDRRRGSDPEMGK -PRRNPOEGVYNALQKDKMAEYSEIGTKGERRRGKHGDLGYQ 144

QY 425 STATKDTYDALHMQALPPR 443  
 DB 145 STATKDTYDALHMQALPPR 163

RESULT 14  
 CD32\_RABIT STANDARD; PRT; 165 AA.  
 AC Q9TUF8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 25-OCT-2004 (Rel. 45, Last annotation update)  
 DE T-cell surface glycoprotein CD3 zeta chain precursor (T-cell receptor T3 zeta chain).  
 GN Name=CD32;  
 OS Oryctolagus cuniculus (Rabbit).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Lagomorpha; Leporidae; Oryctolagus.  
 OX NCBI\_TaxID=9986;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=B/J x Chbb.HM;  
 RA Isono T., Nishimura M.;  
 RT "Rabbit CD3 zeta.";  
 RL Submitted (Nov-1999) to the EMBL/GenBank/DBJ databases.  
 CC -I- FUNCTION: Probable role in assembly and expression of the TCR complex as well as signal transduction upon antigen triggering.  
 CC -I- SUBUNIT: The TCR/CD3 complex of T lymphocytes consists of either a TCR alpha/beta or TCR gamma/delta heterodimer coexpressed at the cell surface with the invariant subunits of CD3 labeled gamma, delta, epsilon, zeta, and eta (By similarity).  
 CC -I- SUBCELLULAR LOCATION: Type I membrane protein.  
 CC -I- PTM: Phosphorylated on Tyr residues after T-cell receptor triggering (By similarity).  
 CC -I- SIMILARITY: Belongs to the CD32/FCER1G family.  
 CC -I- SIMILARITY: Contains 3 ITAM domains.

-----

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation -





**THIS PAGE BLANK (USPTO)**